

Result No.	Score	Query Match	Length	DB	ID	Description
1	1447.4	96.5	3607	11	AK004753	Mus muscu
2	687.4	45.8	845	10	BI688333	603316655
3	648	43.2	810	10	BI657849	603284656
4	589.4	39.3	1094	10	BGI17413	602334132
5	588.2	39.2	662	10	BI694566	603348371
6	517.8	34.5	642	10	BG976005	602846032
7	491.4	32.8	666	10	BI158143	602920783
8	488	32.5	570	10	BI653667	603300323
9	473.4	31.6	768	10	BI646945	603278705
10	465.4	31.0	964	10	BF539781	602049993
11	431.2	28.7	457	9	AW763229	ur-55c07_y
12	423.6	28.2	712	10	BG914492	602813234
13	406.4	27.1	440	9	B8862353	B8862353
14	380	25.3	520	10	BF046412	BF250011A
15	378	25.2	574	10	BF658605	maa98e08
16	353.2	23.5	1370	11	AK015047	Mus muscu
17	315.4	21.0	493	9	AW465562	BP230019B



```
QY 865 TCCGGATGCTGTGTCACCTGAACCTCTACATTCCTTACAAGTCTTTGAGATCAGG 924
Db 1084 TCCGGATGCTGTGTCACCTGAACCTCTACATTCCTTACAAGTCTTTGAGATCAGG 1143
QY 925 TCTGAGAAGAGTGGCAGGATATCAGCATGCGCATGAAGACCATGGGGAGACATC 984
Db 1144 TCTGAGAAGAGTGGCAGGATATCAGCATGCGCATGAAGACCATGGGGAGACATC 1203
QY 985 CTGCCCCACATCAGCAGGAGTGCATTCCTCTTCTGTCATGACGCTGGGATCAAGTCTTT 1044
Db 1204 CTGCCCCACATCAGCAGGAGTGCATTCCTCTTCTGTCATGACGCTGGGATCAAGTCTTT 1263
QY 1045 CAAGACAACTTCGGGTGGAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCCCTGGTG 1104
Db 1264 CAAGACAACTTCGGGTGGAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCCCTGGTG 1323
QY 1105 TACAAGGCCAGTCCCGAGAGTTTCACTATGAGAGCGGGAACCTGTCGGCCGCTACATT 1164
Db 1324 TACAAGGCCAGTCCCGAGAGTTTCACTATGAGAGCGGGAACCTGTCGGCCGCTACATT 1383
QY 1165 CCATTCGGAGAGGGGATTTTACTACCGCGGCCCATTTTGGAGAGCGCCCTACTCAC 1224
Db 1384 CCATTCGGAGAGGGGATTTTACTACCGCGGCCCATTTTGGAGAGCGCCCTACTCAC 1443
QY 1225 ATTCTCAACCTCACCAGGAGTCTTTAAGGGATCTCCAGGACAGAACAATGACATA 1284
Db 1444 ATTCTCAACCTCACCAGGAGTCTTTAAGGGATCTCCAGGACAGAACAATGACATA 1503
QY 1285 GAAGCCAGTGGCATGATGAGAGCCACCTCAACAATACTTCTCTTTTCAACAACCCACT 1344
Db 1504 GAAGCCAGTGGCATGATGAGAGCCACCTCAACAATACTTCTCTTTTCAACAACCCACT 1563
QY 1345 AAAATCCTATCTCCAGAGTATTCCTGGGACTATCAGATAGGCCCTGCTTCAGATATATA 1404
Db 1564 AAAATCCTATCTCCAGAGTATTCCTGGGACTATCAGATAGGCCCTGCTTCAGATATATA 1623
QY 1405 AGTGTCAAGTAGCTGGCAGACAAAAGAGTATAATTTGGTTAGAATAATGCTGACTT 1464
Db 1624 AGTGTCAAGTAGCTGGCAGACAAAAGAGTATAATTTGGTTAGAATAATGCTGACTT 1683
QY 1465 CAAATTGTGATGAAACTTGACACTATTTCT 1495
Db 1684 CAAATTGTGATGAAACTTGACACTATTTACT 1714

RESULT 2
LOCUS BI688333
DEFINITION 845 bp mRNA linear EST 18-SEP-2001
603316655F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5356621 5',
mRNA sequence.
ACCESSION BI688333
VERSION BI688333.1 GI:15650962
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 845)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1906 row: n column: 14
High quality sequence stop: 704.
Location/Qualifiers
```

```
source 1..845
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5356621"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally; Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 222 a 233 c 212 g 178 t
ORIGIN

Query Match 45.8%; Score 687.4; DB 10; Length 845;
Best Local Similarity 98.9%; Pred. No. 1,le-186;
Matches 713; Conservative 0; Mismatches 6; Indels 2; Caps 2;

QY 37 CGATTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGGAGTT-GGAACCCCTGTACCTTTCCT 95
Db 1 CGATTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGGAGTGGGAGCCCTGTACCTTTCCT 60
QY 96 TTCTCTGCTGAGCCCTGCTCTTAGGCGAGGCGAGAGCTCGACAGAACTCGGTTCGCTTT 155
Db 61 TTCTCTGCTGAGCCCTGCTCTTAGGCGAGGCGAGAGCTCGACAGAACTCGGTTCGCTTT 120
QY 156 GCTGTTTGTGGAGGGAACACAGCTGAGGATGAGGCTGACTTTGAACTCAAGATCT 215
Db 121 GCTGTTTGTGGAGGGAACACAGCTGAGGATGAGGCTGACTTTGAACTCAAGATCT 180
QY 216 GCTTACCCAGTCTCTTGGAAATTAAGGCGCTGTACTACATTTGCTCGGACCTAAGATTTT 275
Db 181 GCTTACCCAGTCTCTTGGAAATTAAGGCGCTGTACTACATTTGCTCGGAGCTAAGATTTT 240
QY 276 CATGATCACTATGCTTCAAGATCTCCATGTCAACAAGATCTCCATGTCCAAGATCCAAGTC 335
Db 241 CATGATCACTATGCTTCAAGATCTCCATGTCAACAAGATCTCCATGTCCAAGATCCAAGTC 300
QY 336 AGAACAAGTCTTCCATCTCAAGATCTGATGATGATGATGATGATGATGATGATGATGAT 395
Db 301 AGAACAAGTCTTCCATCTCAAGATCTGATGATGATGATGATGATGATGATGATGATGAT 360
QY 396 AAAAGTAATCTCTTGTGATGCTGATTTCTCAACCGCTGTTGCTGTTTGGGAATATGT 455
Db 361 AAAAGTAATCTCTTGTGATGCTGATTTCTCAACCGCTGTTGCTGTTTGGGAATATGT 420
QY 456 CAACAGAAATTCAGAGGTTGGTGAGAACAGATGGCAGAGGACTGGTGGTTCCTCAAGCTG 515
Db 421 CAACAGAAATTCAGAGGTTGGTGAGAACAGATGGCAGAGGACTGGTGGTTCCTCAAGCTG 480
QY 516 GTTTAAATATGGACCCACACAGTTATCAAGAAGACACACCTAGAAGCGGAGAGAAAAGG 575
Db 481 GTTTAAATATGGACCCACACAGTTATCAAGAAGACACACCTAGAAGCGGAGAGAAAAGG 540
QY 576 TAGAAATGGAGATCGCATTTCAAGAGCTCAGCTATGAGGCTGTTCAATCCAAGAACC 635
Db 541 TAGAAATGGAGATCGCATTTCAAGAGCTCAGCTATGAGGCTGTTCAATCCAAGAACC 600
QY 636 CCCGGATGTTTTCACAGTGTACCCCGTGGAAAGGCCGCGATTTGTGTGGGAAGGCACTTATGA 695
Db 601 CCCGGATGTTTTCACAGTGTACCCCGTGGAAAGGCCGCGATTTGTGTGGGAAGGCACTTATGA 660
QY 696 CACAGCTCTCTGGAA-AAGTACTACCCACACAGAAACTCACTGTGGGCTGACAGTGT 754
Db 661 CACAGCTCTCTGGAAACAAAGTACTACCCACACAGAAACTCACTGTGGGCTGACAGTGT 720
QY 755 T 755
Db 721 T 721
```

```

RESULT 3
BI657849 810 bp mRNA linear EST 12-SEP-2001
LOCUS 603284656F1 NIH_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5329021 5',
DEFINITION mRNA sequence.
ACCESSION BI657849
VERSION 1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 810)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1834 row: p column: 14
High quality sequence stop: 735.
FEATURES
source
1. 810
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5329021"
/clone_lib="NIH_CGAP_Mam4"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert 2.5 kb. Library constructed by Life
Technologies, catalog # 12018-016. Investigators providing
samples: Lothar Hennighausen/Priscilla Furth, NIH
Reference for transgenic model: Li et al., Cell Growth and
Differentiation 7, 3-11 (1996). Note: this is a NIH_CGAP
Library."
BASE COUNT 222 a 199 c 198 g 191 t
ORIGIN
Query Match 43.2%; Score 648; DB 10; Length 810;
Best Local Similarity 91.6%; Pred. No. 2.4e-175;
Matches 753; Conservative 0; Mismatches 35; Indels 34; Gaps 5;
Qy 268 AAGATTTTCATGATCACTATGCTTCAAGATCTCCATGTCAACAGATCTCCATGTCAAGA 327
Dy 1 AAGATTTTCATGATCACTATGCTTCAAGATCTCCATGTCAACAGATCTCCATGTCAAGA 60
Qy 328 TCCAAGTCAGAAACAAGTCTTCCATCTCAAGATCTGGATCACAGGAGAAATAATGAAT 387
Dy 61 TCCAAGTCAGAAACAAGTCTTCCATCTCAAGATCTGGATCACAGGAGAAATAATGAAT 120
Qy 388 GTCAAGGAAAAGTAATCTCTTGATGCTGATGTCTCAACCGGTGTGCTGTTTTGG 447
Dy 121 GTCAAGGAAAAGTAATCTCTTGATGCTGATGTCTCAACCGGTGTGCTGTTTTGG 180
Qy 448 GAATATGTCAACAGATTTCCAGAGTTGGTGAGAACAGATGGCAAGAGGACTGGTGTTC 507
Dy 181 GAATATGTCAACAG-----CCAGAGAGGCTCTTTC 210
Qy 508 CCAAGCTGTTTAAATAATGGGACCCACAGTTTATCAAGAAGACAACGTAGAGGACGGAGA 567
Dy 211 TTGTGGATATATCACAAAGACCCACAGTTTATCAAGAAGACAACGTAGAGGACGGAGA 270

```



```
/dev stage="10 months. virgin"
```

```
dev_stage= IO 1100
lab_host="192.168.1.100"
```

```

/abn_nost="DHIUB"
/mam="mammary; Vector: pCMV-sport6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

```

providing samples: street	288 c	291 a	190 f
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			
61			
62			
63			
64			
65			
66			
67			
68			
69			
70			
71			
72			
73			
74			
75			
76			
77			
78			
79			
80			
81			
82			
83			
84			
85			
86			
87			
88			
89			
90			
91			
92			
93			
94			
95			
96			
97			
98			
99			
100			

## ORIGIN

Query Match 39.38; Score 589.4; DB 10; Length 1094;

Best Local Similarity	96.28;	Pred. No. 2.2e-158;	Indels	3;	Gaps	1;
Matches	615;	Conservative	0;	Mismatches	21;	
Matches	615;	Conservative	0;	Mismatches	21;	

QV 523 AATGGGACCCACAGTTATCAAGAAGACAAACGTAGAAGGACGGAGAGAAAGGGTAGAAAT 582

9} 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1

Db 94 AACAGGACCCACAGTTATCAAGAAGACAAACGTAGAAGGACGGAGAGAAAAGGGTAGAAAT 153

2021年12月31日

QY 583 GGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAGAACCGCCCGGAT 642

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Db 154 GGAGATCGCATTTGAAGAGCCTCAGCTATGGACTGGTTCAATCCAAGAACCGCCGGAT 213

[illegible]

QY 643 GTTTTGACAGTGACCCCGTGGAAGCGCCGATTGTGTGGGAAGGCACTATGACACAGCT 702

\_\_\_\_\_

Db 214 GTTTGACAGTGACCCCGTGAAGGCCCGATTGTGTGGAGGCACCTATGACACAGCT 273

.....

QY 703 CTGCTGGAAAGTACTACGCCACACAGAAACTCACTGTGGGGCTGACAGTCTTTGCTGTG 762

\_\_\_\_\_

Db 274 CTGCTGGAAAGTACTACGCCACACAGAAACTCACTGTGGGGCTGACAGTGTTCCTGTG 333

Qy 763 GGAAAGTACATTGAGCATTTACTTAGAAGACTTTCTGGAGTCTGCTGACATGTACTTCATG 822

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Db 334 GGAAGTACATTGAGCATTTACTTAGAAGACTTTCTGGAGTCTGCTGACATGTACTTCATG 393

Qy 823 GTGGCCATCGGTCATATTTACGTCATGATAGACGACACCTCCCGATGCCCTGTCGTG 882

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

## RESULT

RESUL  
AT694566

00C660T9  
LOCUS

LOCUS	BI694566	NC
DEFINITION	603349371E1	NC

DEFINITION 60334837: tRNA GCG

Accession	mRNA seq
U00096.1	U00096.1
U00097.1	U00097.1
U00098.1	U00098.1
U00099.1	U00099.1
U00100.1	U00100.1
U00101.1	U00101.1
U00102.1	U00102.1
U00103.1	U00103.1
U00104.1	U00104.1
U00105.1	U00105.1
U00106.1	U00106.1
U00107.1	U00107.1
U00108.1	U00108.1
U00109.1	U00109.1
U00110.1	U00110.1
U00111.1	U00111.1
U00112.1	U00112.1
U00113.1	U00113.1
U00114.1	U00114.1
U00115.1	U00115.1
U00116.1	U00116.1
U00117.1	U00117.1
U00118.1	U00118.1
U00119.1	U00119.1
U00120.1	U00120.1
U00121.1	U00121.1
U00122.1	U00122.1
U00123.1	U00123.1
U00124.1	U00124.1
U00125.1	U00125.1
U00126.1	U00126.1
U00127.1	U00127.1
U00128.1	U00128.1
U00129.1	U00129.1
U00130.1	U00130.1
U00131.1	U00131.1
U00132.1	U00132.1
U00133.1	U00133.1
U00134.1	U00134.1
U00135.1	U00135.1
U00136.1	U00136.1
U00137.1	U00137.1
U00138.1	U00138.1
U00139.1	U00139.1
U00140.1	U00140.1
U00141.1	U00141.1
U00142.1	U00142.1
U00143.1	U00143.1
U00144.1	U00144.1
U00145.1	U00145.1
U00146.1	U00146.1
U00147.1	U00147.1
U00148.1	U00148.1
U00149.1	U00149.1
U00150.1	U00150.1
U00151.1	U00151.1
U00152.1	U00152.1
U00153.1	U00153.1
U00154.1	U00154.1
U00155.1	U00155.1
U00156.1	U00156.1
U00157.1	U00157.1
U00158.1	U00158.1
U00159.1	U00159.1
U00160.1	U00160.1
U00161.1	U00161.1
U00162.1	U00162.1
U00163.1	U00163.1
U00164.1	U00164.1
U00165.1	U00165.1
U00166.1	U00166.1
U00167.1	U00167.1
U00168.1	U00168.1
U00169.1	U00169.1
U00170.1	U00170.1
U00171.1	U00171.1
U00172.1	U00172.1
U00173.1	U00173.1
U00174.1	U00174.1
U00175.1	U00175.1
U00176.1	U00176.1
U00177.1	U00177.1
U00178.1	U00178.1
U00179.1	U00179.1
U00180.1	U00180.1
U00181.1	U00181.1
U00182.1	U00182.1
U00183.1	U00183.1
U00184.1	U00184.1
U00185.1	U00185.1
U00186.1	U00186.1
U00187.1	U00187.1
U00188.1	U00188.1
U00189.1	U00189.1
U00190.1	U00190.1
U00191.1	U00191.1
U00192.1	U00192.1
U00193.1	U00193.1
U00194.1	U00194.1
U00195.1	U00195.1
U00196.1	U00196.1
U00197.1	U00197.1
U00198.1	U00198.1
U00199.1	U00199.1
U00200.1	U00200.1
U00201.1	U00201.1
U00202.1	U00202.1
U00203.1	U00203.1
U00204.1	U00204.1
U00205.1	U00205.1
U00206.1	U00206.1
U00207.1	U00207.1
U00208.1	U00208.1
U00209.1	U00209.1
U00210.1	U00210.1
U00211.1	U00211.1

ACCESSION	BI694566	DT04566
1	2	3
4	5	6
7	8	9
10	11	12
13	14	15
16	17	18
19	20	21
22	23	24
25	26	27
28	29	30
31	32	33
34	35	36
37	38	39
40	41	42
43	44	45
46	47	48
49	50	51
52	53	54
55	56	57
58	59	60
61	62	63
64	65	66
67	68	69
70	71	72
73	74	75
76	77	78
79	80	81
82	83	84
85	86	87
88	89	90
91	92	93
94	95	96
97	98	99
100	101	102
103	104	105
106	107	108
109	110	111
112	113	114
115	116	117
118	119	120
121	122	123
124	125	126
127	128	129
130	131	132
133	134	135
136	137	138
139	140	141
142	143	144
145	146	147
148	149	150
151	152	153
154	155	156
157	158	159
160	161	162
163	164	165
166	167	168
169	170	171
172	173	174
175	176	177
178	179	180
181	182	183
184	185	186
187	188	189
190	191	192
193	194	195
196	197	198
199	200	201
202	203	204
205	206	207
208	209	210
211	212	213
214	215	216
217	218	219
220	221	222
223	224	225
226	227	228
229	230	231
232	233	234
235	236	237
238	239	240
241	242	243
244	245	246
247	248	249
250	251	252
253	254	255
256	257	258
259	260	261
262	263	264
265	266	267
268	269	270
271	272	273
274	275	276
277	278	279
280	281	282
283	284	285
286	287	288
289	290	291
292	293	294
295	296	297
298	299	300
301	302	303
304	305	306
307	308	309
310	311	312
313	314	315
316	317	318
319	320	321
322	323	324
325	326	327
328	329	330
331	332	333
334	335	336
337	338	339
340	341	342
343	344	345
346	347	348
349	350	351
352	353	354
355	356	357
358	359	360
361	362	363
364	365	366

VERSION BI694566.1

**KEYWORDS** EST.

**SOURCE** house mouse.

**ORGANISM** *Mus musculus*

**Eukaryota: M**

Mammalia; Eu

```
Db 538 TGAGATCAGGTCTCAGAAAGAGGTGGCAGGATATACAGCATGATCGCATGAAGACCATTTGG 597
Qy 975 GGAGCACATCTGCGCCACATCCAGCAGCGAGTGCAGTCTCTC-TTCTGTCATGGACGTGG 1033
Db 598 GGAGCACATCTGCGCCACATCCAGCAGCGAGTGCAGTCTCTC-TTCTGTCATGGACGTGG 657
Qy 1034 ATCAA 1038
Db 658 ATCAA 662

RESULT 6
Bg976005
LOCUS 602846032F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4976863 5',
DEFINITION mRNA sequence.
ACCESSION Bg976005
VERSION Bg976005.1 GI:14363642
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
1 (bases 1 to 642)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMI0970 row: o column: 08
High quality sequence start: 3
High quality sequence stop: 642.
FEATURES
source
1. .642
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4976863"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 160 a 147 c 199 g 136 t
ORIGIN

Query Match 34.5%; Score 517.8; DB 10; Length 642;
Best Local Similarity 97.3%; Pred. No. 7e-138;
Matches 571; Conservative 0; Mismatches 7; Indels 9; Gaps 4;

Qy 543 AGAAGACAACGTAGAGGACGAGAGAGAAAGGGGTAGAAATCGAGATCGCATTCGAAGAGCC 602
Db 60 AGAAGACAACGTAGAGGACGAGAGAGAAAGGGGTAGAAATCGAGATCGCATTCGAAGAGCC 119
Qy 603 TCAGCTATGGAGTGGTTCAATCCAAAGAACCGCCCGGATGTTTTCAGATGACGACCCCGTG 662
Db 120 TCAGCTATGGAGTGGTTCAATCCAAAGAACCGCCCGGATGTTTTCAGATGACGACCCCGTG 179
Qy 663 GAAGGCCCGCATTTGTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAGAAAGTACTACGC 722
Db 180 GAAGGCCCGCATTTGTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAGAAAGTACTACGC 239
```

```
Qy 723 CACAGAAAACCTACTGTGGGGCTGACAGTC-TTTGCTGTGGAAAGTACATTTGAGCATTT 781
Db 240 CACAGAAAACCTACTGTGGGGCTGACAGTC-TTTGCTGTGGAAAGTACATTTGAGCATTT 299
Qy 782 ACTTAGAAGACTTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTGCATAT 841
Db 300 ACTTAGAAGACTTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTGCATAT 359
Qy 842 TTTACGTCATGATA--GAGCACACCTCCCGGATGCGCTGCTGCGACCTGAACCTCTAC 898
Db 360 TTTACGTCATGGAATAGAAGACACCTCCCGGATGCGCTGCTGCGACCTGAACCTCTAC 419
Qy 899 ATTCTTTTACAAGTCTTTTTCAGATCAGAGTCTGAGAGAGAGTGGCAGGATATCAGCATGATGC 958
Db 420 ATTCTTTTACAAGTCTTTTTCAGATCAGAGTCTGAGAGAGAGTGGCAGGATATCAGCATGATGC 479
Qy 959 GCATGAAGACCATTTGGGAGGACACATCTCTGGCCACATCCAGCAGGAGTGCAGTTCCTCT 1018
Db 480 GCATGAAGACCATTTGGGAGGACACATCTCTGGCCACATCCAGCAGGAGTGCAGTTCCTCT 539
Qy 1019 TCTGTCATGGAGCTGGATCAAGTCTTTCAAGACAACTTTCGGGGTGAAGTCTCTGGCCAGC 1078
Db 540 TCTGTCATGGAGCTGGATCAAGTCTTTCAAGACAACTTTCGGGGTGAAGTCTCTGGCCAGC 596
Qy 1079 TGTGATGACAGCTCCAGGCGCTGGT-GGTACAAGGCCAGTCCCGAGAA 1124
Db 597 -TGTAGCACAGCTCCAGGCGCTGGTGGGTACAAAGGACAGTCCCGAGAA 642

RESULT 7
Bg976005
LOCUS 602920783F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5061116 5',
DEFINITION mRNA sequence.
ACCESSION Bg976005
VERSION Bg976005.1 GI:14618144
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
1 (bases 1 to 666)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMI1165 row: e column: 21
High quality sequence stop: 665.
FEATURES
source
1. .666
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5061116"
/clone_lib="NIH_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert 2 kb. Library constructed by Life
Technologies, catalog #12017-018. Investigators providing
samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference
for transgenic model: Xu et al., Nature Genetics 22, 37-43
(1999). Note: this is a NCI_CGAP Library."
BASE COUNT 185 a 158 c 160 g 163 t
```

ORIGIN

Query Match 32.8%; Score 491.4; DB 10; Length 666;  
Best Local Similarity 90.8%; Pred. No. 2.9e-130;  
Matches 595; Conservative 0; Mismatches 16; Indels 44; Gaps 5;  
QY 36 CCGATTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGAGTTGGAACCCCTGTACCTTCCT 95  
DB 1 CCGATTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGAGTTGGAACCCCTGTACCTTCCT 60  
QY 96 TTCTCTCTGCTGACCCCTGCTCTTAGGAGGCCAGCTCGACAGCACTCGGTTGCTTTT 155  
DB 61 TTCTCTCTGCTGACCCCTGCTCTTAGGAGGCCAGCTCGACAGCACTCGGTTGCTTTT 120  
QY 156 GCTGTTTGTCTTGGAGGAACACAGCTGAGCTGAGGCTGACTTTGAACCTCAAGAGATCT 215  
DB 121 GCTGTTGCTTGGAGGAACACAGCTGAGCTGAGGCTGACTTTGAACCTCAAGAGATCT 180  
QY 216 GCTTACCCCTGCTCTGGAATTAAGGCTGTACTACATTTGCTGCTGAGCTCAAGATTTT 275  
DB 181 GCTTACCCCTGCTCTGGAATTAAGGCTGTACTACATTTGCTGCTGAGCTCAAGATTTT 240  
QY 276 CATGATCACTATGCTTCAAGATCTCCATGTCACAGATCTCCATGTCAGATCCAAAGTC 335  
DB 241 CATGATCACTATGCTTCAAGATCTCCATGTCACAGATCTCCATGTCAGATCCAAAGTC 300  
QY 336 AGAAACAAGTCTTCCATCTCCTCAAGATCTGGATCACAGGAGAAAATTAATGAATGTCAGGG 395  
DB 301 AGAAACAAGTCTTCCATCTCCTCAAGATCTGGATCACAGGAGAAAATTAATGAATGTCAGGG 360  
QY 396 AAAAGTAATCTGCTGATGCTGATGCTCAACCGTGGTGTGCTGCTGTTTGGGAATATGT 455  
DB 361 AAAAGTAATCTGCTGATGCTGATGCTCAACCGTGGTGTGCTGCTGTTTGGGAATATGT 420  
QY 456 CAACAG-----AATTCAGAGGTTGGTGA 479  
DB 421 CAACAGCCCAAGAGGCTCTTCTTGTGGATATATACACAAAAATTCAGAGGTTGGTGA 480  
QY 480 GAACAGATGCT--AGAAAGACTGGTGTCCCA--AGCTGCTTTAAA--AATGGGACCCACA 535  
DB 481 GAACAGATGCCAGAAAGACTGGTGTCCCAAGCTGTTTAAACAATGGGACCCACA 540  
QY 536 GTTATCAAGAGAACAGCTAGAGGACGGA---GAGAAAGGTTAGAAATGAGATCGC 591  
DB 541 GTTATCAAGAGAACAGCTAGAGGACGGAAGGCTGAGACAAAGGCTCAGAAAATGAGATCGC 600  
QY 592 ATTGAAGAGCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGGGATGTTT 646  
DB 601 ATTGAAGAGCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGGGATGTTT 655

RESULT 8

BI653667 570 bp mRNA linear EST 12-SEP-2001  
LOCUS 603300333F1 NTH\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:5340868 5',  
DEFINITION mRNA sequence.  
ACCESSION BI653667  
VERSION BI653667.1 GI:15567903  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 570)  
NIH-MGC <http://mgi.nhl.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLML1865 row: n column: 05  
High quality sequence stop: 570.  
FEATURES  
Location/Qualifiers  
1..570  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5340868"  
/clone\_lib="NIH\_CGAP\_Mam3"  
/tissue\_type="tumor, gross tissue"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert 2 kb. Library constructed by Life  
Technologies, catalog #12017-018. Investigators providing  
samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference  
for transgenic model: Xu et al., Nature Genetics 22, 37-43  
(1999). Note: this is a NCI\_CGAP Library."  
BASE COUNT 162 a. 115 c 156 g 137 t  
ORIGIN  
Query Match 32.5%; Score 488; DB 10; Length 570;  
Best Local Similarity 93.7%; Pred. No. 2.5e-129;  
Matches 534; Conservative 0; Mismatches 0; Indels 36; Gaps 1;  
QY 362 CTGGATCACAGGAGAAAATAATGAATGTCAGGAGAAAAGTAATCCTGTGTGATGCTGATTC 421  
DB 1 CTGGATCACAGGAGAAAATAATGAATGTCAGGAGAAAAGTAATCCTGTGTGATGCTGATTC 60  
QY 422 TCTCAACCGTGGTGTGCTGCTTTGGGAATATGTCACACCCAGAGGCTCTTCTTGT 120  
DB 61 TCTCAACCGTGGTGTGCTGCTTTGGGAATATGTCACACCCAGAGGCTCTTCTTGT 120  
QY 462 -----AATTCAGAGGTTGGTGTGAGAACAGATGGCAGAGGACTGGTGGT 505  
DB 121 GGATATATCACAAAAATTCAGAGGTTGGTGTGAGAACAGATGGCAGAGGACTGGTGGT 180  
QY 506 TCCCAAGCTGGTTTAAAAATGGGACCCACAGTTATCAAGAAGACACAGTATAGAAGGACGGA 565  
DB 181 TCCCAAGCTGGTTTAAAAATGGGACCCACAGTTATCAAGAAGACACAGTATAGAAGGACGGA 240  
QY 566 GAGAAAAGGTAGAAATGGAGATCGCATTCAGACAGCTCAGCTATGGAGCTGGTTCATC 625  
DB 241 GAGAAAAGGTAGAAATGGAGATCGCATTCAGACAGCTCAGCTATGGAGCTGGTTCATC 300  
QY 626 CAAAGAACCCCGCGGATGTTTTCACAGTGCACCCGTTGGAAGGCCGCGATTTGTGTGGGAAG 685  
DB 301 CAAAGAACCCCGCGGATGTTTTCACAGTGCACCCGTTGGAAGGCCGCGATTTGTGTGGGAAG 360  
QY 686 GCACTTATGACACAGCTGCTGCTGAAAAAGTACTACGCCACACAGAAACTCAGTGTGGGGC 745  
DB 361 GCACTTATGACACAGCTGCTGCTGAAAAAGTACTACGCCACACAGAAACTCAGTGTGGGGC 420  
QY 746 TGACAGTGTGCTGTGGGAAAGTACATTCAGCATTTACTTAGAAGACTTCTCGAGTCTG 805  
DB 421 TGACAGTGTGCTGTGGGAAAGTACATTCAGCATTTACTTAGAAGACTTCTCGAGTCTG 480  
QY 806 CTGACATGTACTTTCATGTTGGCCATCGGCTCATATTTTACGTGTCATGATAGACGACCT 865  
DB 481 CTGACATGTACTTTCATGTTGGCCATCGGCTCATATTTTACGTGTCATGATAGACGACCT 540  
QY 866 CCCGGATGCTGTGTCACCTGAACCCCTC 895  
DB 541 CCCGGATGCTGTGTCACCTGAACCCCTC 570  
RESULT 9  
BI646945 768 bp mRNA linear EST 12-SEP-2001  
LOCUS BI646945

[illegible]

QY 153 TTTGCTGTTTCTTGGAGGAAACACAGCTGACGATGAGGCTGACTTTGAACCTCAAGAGA 212  
 Db 121 T-----GCTGTTGCTTGGAGGAAACACAGCTGACGATGAGGCTGAC-TTGAACCTCAAGAGA 175  
 QY 213 TCTGCTTACCCAGTCTCTCTGGAAATTAAGGCTGTACTACATCTTCCCTGGACCTAAGAT 272  
 Db 176 TCTGCTTACCCAGTCTCTCTGGAAATTAAGGCTGTACTACATCTTCCCTGGACCTAAGAT 235  
 QY 273 TTTTCATGATCAGTATGCTTTCAGATCTCCATGTCACAAAGATCTCCATGTCAGATGCCAA 332  
 Db 236 TTTTCATGATCAGTATGCTTTCAGATCTCCATGTCACAAAGATCTCCATGTCAGATGCCAA 295  
 QY 333 GTCAGAAACAAAGTCTTCCATCTCAAGATCTGGATCACAGGAGAAAATTAATGAATGTCAA 392  
 Db 296 GTCAGAAACAAAGTCTTCCATCTCAAGATCTGGATCACAGGAGAAAATTAATGAATGTCAA 355  
 QY 393 GGGAAAAGTAATCTGTTGATGCTGATGCTCAACCGTGGTGTCTGTTTGGGAATA 452  
 Db 356 GGGAAAAGTAATCTGTTGATGCTGATGCTCAACCGTGGTGTCTGTTTGGGAATA 415  
 QY 453 TGTCAACAGAAATCCAGAGGTTGGTGAGAAACAGATGGCAGAGGACTGGTGGTTCCCAAG 512  
 Db 416 TGTCAACA----- 423  
 QY 513 CTGCTTTAAAAATGGACCCACACAGTTATCAAGAAGACAAACGTAGAAGCGGAGAGAAA- 571  
 Db 424 -----GCACCCACAGTTATCAAGAAGACAAACGTAGAAGCGGAGAGAAAAC 469  
 QY 572 AGGGTAGAATGAGATCGCATTTGAAGAGCCCTCAGCTATGGGACTGGTTCATTCCAAAGA 631  
 Db 470 AGGGTAGAATGAGATCGCATTTGAAGAGCCCTCAGCTATGGGACTGGTTCATTCCAAAGA 529  
 QY 632 ACCGCCCGGATGTTTGACAGTCAACCCGCGGAGGCGGATGTTGGGAAAGGCACCTT 691  
 Db 530 ACCGCC--GGATGTTTGACAGTCAACCCGCGGAGGCGGGA-TGTGTGGGAAAGGCACCTT 586  
 QY 692 ATGACACAGCTCTGCTGGAAAGTACTACGCCACACAGAAACTCAGTGTGGGCTGACAG 751  
 Db 587 ATGACACAGCTCTGCTGGAAAGTCTTACGCCACACAGAAACTCAGTGTGGGCGCTGAC 646  
 QY 752 TGTCTGCTGGGAAAGTACATTTAGCATTTTACGATGATAGACAGACCTCCCGG 811  
 Db 647 AAGGGTGTCTGGGGAAACGACCTGGGACCTTAATTTAAACACTCGGGGAACTCGCGGACAG 706  
 QY 812 TGTACTTCATG-CTTGGCCATCGGCTCATATTTTACGTCATGATAGACAGACCTCCCGG 870  
 Db 707 TGTCTCCGGTGTGTGGGCCAAGGGGCCCAATACAGCGTGTATAGACAAAACCCCGG 766  
 QY 871 ATGCTGTCTC 879  
 Db 767 GAGTAAGTC 775  
 RESULT 11  
 AW762229 457 bp mRNA linear EST 04-MAY-2000  
 LOCUS ur55c07.y1 NCI\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:3154188 5'  
 DEFINITION similar to gb:M26925 Mouse galactosyltransferase mRNA, complete cds  
 (MOUSE);, mRNA sequence.  
 ACCESSION AW762229  
 VERSION AW762229.1 GI:7694150  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 457)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Other\_ESTs: ur55c07.x1

Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
 cDNA Library preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 image.llnl.gov/image/html/iresources.shtml

MGI:1056944  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 399.  
 Location/Qualifiers  
 1. .457  
 /organism="Mus musculus"  
 /strain="129 - C57/B6 - FVB"  
 /db\_xref="taxon:10090"  
 /clone\_lib="IMAGE:3154188"  
 /clone\_lib="NCI\_CGAP\_Mam3"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="10 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: Salt;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH  
 Reference for transgenic model: Xu et al., Nature Genetics  
 22, 37-43 (1999)."

BASE COUNT 112 a 116 c 106 g 123 t  
 ORIGIN

Query Match 28.7%; Score 431.2; DB 9; Length 457;  
 Best Local Similarity 99.3%; Pred. No. 5.2e-113;  
 Matches 433; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 26 ATGAGAAGTACCGATTTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGAGCTTGGAAACCTT 85  
 Db 22 ATCAGAAGTACCGATTTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGAGCTTGGAAACCTT 81  
 QY 86 GTACCTTCTTCTGCTGAGCCCTGCTCTAGGAGGCGCAGAGCTCGACAGAACT 145  
 Db 82 GTACCTTCTTCTGCTGAGCCCTGCTCTAGGAGGCGCAGAGCTCGACAGAACT 141  
 QY 146 CGCTTCTTCTGCTTGTGTTGGAGGAACACAGCTGAGCATGAGGCTGACCTTTTGAACCT 205  
 Db 142 CGTCTTCTTCTGCTTGTGTTGGAGGAACACAGCTGAGCATGAGGCTGACCTTTTGAACCT 201  
 QY 206 CAAGAGATCTGCTTACCCAGTCTCCTTGGAAATTAAGGCCCTGTACTACATTTGCCCTGGAC 265  
 Db 202 CAAGAGATCTGCTTACCCAGTCTCCTTGGAAATTAAGGCCCTGTACTACATTTGCCCTGGAC 261  
 QY 266 CTAAGATTTTCATGATCAGTACTATGCTTCAAGATCTCCATGTCACAAAGATCTCCATGTCAA 325  
 Db 262 CTAAGATTTTCATGATCAGTACTATGCTTCAAGATCTCCATGTCACAAAGATCTCCATGTCAA 321  
 QY 326 GATCCAAAGTCAGAAACAAAGCTTCCATCTCCTCAAGATCTGGATCACAGGAGAAAATTAATGA 385  
 Db 322 GATCCAAAGTCAGAAACAAAGCTTCCATCTCCTCAAGATCTGGATCACAGGAGATTAATGA 381  
 QY 386 ATGTCAAGGAGAAAGTAATCTTGTGATGCTGATGTTCTCAACCGTGGTGTCTGCTGTTT 445  
 Db 382 ATGTCAAGGAGAAAGTAATCTTGTGATGCTGATGTTCTCAACCGTGGTGTCTGCTGTTT 441  
 QY 446 GGAATATGTCAACAG 461  
 Db 442 GGAATATGTCAACAG 457  
 RESULT 12  
 BG914492  
 LOCUS BG914492 linear EST 05-JUN-2001









GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 17:58:51 : Search time 3258.53 seconds  
(without alignments)  
9959.199 Million cell updates/sec

Title: US-09-863-475A-3  
Perfect score: 1500  
Sequence: 1 CCTCCCTGTGAGACTCTTC.....CTTGACACTATTTCTAACCA 1500

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/2/pna/PCTUS\_COMB.seq:\*

2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq:\*

3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq:\*

4: /cgn2\_6/ptodata/2/pna/US080\_COMB.seq:\*

5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq:\*

6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq:\*

7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq:\*

8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq:\*

9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq:\*

10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq:\*

11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq:\*

12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq:\*

13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq:\*

14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq:\*

15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq:\*

16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq:\*

17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:\*

18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:\*

19: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq:\*

20: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq:\*

21: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq:\*

22: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq:\*

23: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq:\*

24: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq:\*

25: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq:\*

26: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq:\*

27: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:\*

28: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq:\*

29: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq:\*

30: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq:\*

31: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq:\*

32: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq:\*

33: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:\*

34: /cgn2\_6/ptodata/2/pna/US099A\_COMB.seq:\*

35: /cgn2\_6/ptodata/2/pna/US099B\_COMB.seq:\*

36: /cgn2\_6/ptodata/2/pna/US099C\_COMB.seq:\*

37: /cgn2\_6/ptodata/2/pna/US100\_COMB.seq:\*

38: /cgn2\_6/ptodata/2/pna/US101\_COMB.seq:\*

39: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq:\*

40: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq:\*

41: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq:\*

42: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq:\*

43: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq:\*

44: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq:\*

45: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq:\*

46: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq:\*

47: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq:\*

48: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq:\*

49: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq:\*

50: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq:\*

51: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq:\*

52: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq:\*

53: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq:\*

54: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq:\*

55: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq:\*

56: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq:\*

57: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq:\*

58: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq:\*

59: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq:\*

60: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq:\*

61: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq:\*

62: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq:\*

63: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq:\*

64: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq:\*

65: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq:\*

66: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq:\*

67: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq:\*

68: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq:\*

69: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq:\*

70: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq:\*

71: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq:\*

72: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq:\*

73: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq:\*

74: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq:\*

75: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1500	100.0	1500	7	US-08-379-040-1
2	1500	100.0	1500	12	US-08-823-489-3
3	1500	100.0	1500	22	US-09-593-316-9
4	1500	100.0	1500	33	US-09-863-475-3
5	1500	100.0	1500	33	US-09-863-475A-3
6	1500	100.0	1500	35	US-09-946-034-1
7	1447.4	96.5	3455	75	US-60-360-207-12027
8	1408.4	93.9	3450	5	US-08-188-607B-9
9	1408.4	93.9	3450	7	US-08-378-617-9
10	1408.4	93.9	3450	13	US-08-984-900-9
11	1005.8	67.1	1131	6	US-08-213-200A-3
12	1005.8	67.1	1131	15	US-09-173-270-3
13	726.6	48.4	1423	6	US-08-214-580-2
14	726.6	48.4	1423	6	US-08-260-201-5
15	726.6	48.4	1423	6	US-08-278-282-5
16	726.6	48.4	1423	15	US-09-137-517-2
17	705	47.0	1091	16	US-09-230-091-3
18	702.8	46.9	1140	15	US-09-173-270-1
19	693.8	46.3	1131	6	US-08-213-200A-1
20	693.8	46.3	1131	22	US-09-593-316-5
21	693.8	46.3	1131	36	US-09-994-427A-14
22	693.8	46.3	1131	36	US-09-995-419A-8
23	692.2	46.1	1131	36	US-09-994-427A-16
24	684.2	45.6	1269	6	US-08-328-933A-1
25	684.2	45.6	1269	13	US-08-929-940-1
26	684.2	45.6	1269	22	US-09-593-316-7
27	684.2	45.6	1412	13	US-08-984-900-7
28	679.4	45.3	1412	5	US-08-188-607B-7
29	679.4	45.3	1412	7	US-08-378-617-7
30	663.6	44.2	1127	16	US-09-230-091-1
31	645.6	43.0	1303	22	US-09-593-316-1

32	645.6	43.0	1303	36	US-09-994-427A-17	Sequence 17, Appl
33	645.6	43.0	1303	36	US-09-995-419A-6	Sequence 6, Appl
34	642.4	42.8	1617	22	US-09-593-315-3	Sequence 3, Appl
35	642.4	42.8	1828	5	US-08-188-607B-8	Sequence 8, Appl
36	642.4	42.8	1828	7	US-08-378-617-8	Sequence 8, Appl
37	642.4	42.8	1828	13	US-08-984-900-8	Sequence 8, Appl
38	631.8	42.1	1064	16	US-09-230-091-5	Sequence 5, Appl
39	616.2	41.1	1028	16	US-09-230-091-7	Sequence 7, Appl
40	583.2	38.9	1353	6	US-08-214-580-1	Sequence 1, Appl
41	583.2	38.9	1353	15	US-09-137-517-1	Sequence 1, Appl
42	561.8	37.5	1885	22	US-09-593-316-11	Sequence 11, Appl
43	476.4	31.8	631	17	US-09-371-508-2528	Sequence 2528, Ap
44	476.4	31.8	631	17	US-09-371-508-2528	Sequence 2528, Ap
45	476.4	31.8	631	29	US-09-747-508-2528	Sequence 2528, Ap

## ALIGNMENTS

```

RESULT 1
US-08-379-040-1
: Sequence 1, Application US/08379040
: GENERAL INFORMATION:
: APPLICANT: Cooper, David K.C.
: APPLICANT: Koren, Eugen
: TITLE OF INVENTION: GENETICALLY ENGINEERED ANIMALS FOR USE
: TITLE OF INVENTION: AS ORGAN DONORS
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Patrea L. Pabst
: STREET: 1100 Peachtree Street, Suite 2800
: CITY: Atlanta
: STATE: GA
: COUNTRY: USA
: ZIP: 30309-4530
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/379,040
: FILING DATE:
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/049,817
: FILING DATE: 20-APR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Pabst, Patrea L.
: REGISTRATION NUMBER: 31,284
: REFERENCE/DOCKET NUMBER: BMC100
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (404)-815-6508
: TELEFAX: (404)-815-6555
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1500 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Mus musculus
: PUBLICATION INFORMATION:
: AUTHORS: Larsen, Robert D.
: AUTHORS: Rajan, Valanila P.
: AUTHORS: Ruff, Melissa M.
: AUTHORS: Kukowska-Latallo, Jolanta
: AUTHORS: Cummings, Richard D.
: AUTHORS: Lowe, John B.
: TITLE: Isolation of a cDNA encoding a murine

```

Db 841 TTTTACGTCATGATAGACACACCTCCCGGATGCCTGTGCTGCACCTGAACCCCTCTACAT 900  
Qy 901 TCCTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGCTGCAGGATATCAGCATGATGCGC 960  
Db 901 TCCTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGCTGCAGGATATCAGCATGATGCGC 960  
Qy 961 ATGAAGACCATTTGGGAGACATCTCTGGCCCATCCAGACAGAGTCCGACTTCTCTTC 1020  
Db 961 ATGAAGACCATTTGGGAGACATCTCTGGCCCATCCAGACAGAGTCCGACTTCTCTTC 1020  
Qy 1021 TGCATGGAGCTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAACTCTGGCCAGCTG 1080  
Db 1021 TGCATGGAGCTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAACTCTGGCCAGCTG 1080  
Qy 1081 GTAGCACAGCTCCAGGCTTGGTGTACAAGGCCAGTCCCGAGAGGTTTCACTATGAGAG 1140  
Db 1081 GTAGCACAGCTCCAGGCTTGGTGTACAAGGCCAGTCCCGAGAGGTTTCACTATGAGAG 1140  
Qy 1141 CGGGAACCTCCGCGCGGTACATTCATTCCGAGAGGGGGATTTTACTACACGCGGC 1200  
Db 1141 CGGGAACCTCCGCGCGGTACATTCATTCCGAGAGGGGGATTTTACTACACGCGGC 1200  
Qy 1201 ATTTTGGAGGACGCTACTACATCTCAACCTCACAGGAGTCTTAAAGGGATC 1260  
Db 1201 ATTTTGGAGGACGCTACTACATCTCAACCTCACAGGAGTCTTAAAGGGATC 1260  
Qy 1261 CTCAGGACAAGAACATCACATAGAACGCCAGTGCATGATGAGAGCCACCTCAACAAA 1320  
Db 1261 CTCAGGACAAGAACATCACATAGAACGCCAGTGCATGATGAGAGCCACCTCAACAAA 1320  
Qy 1321 TACTTCTTTTCAACAAACCCACTAAATCTATCTCCAGAGTATTGCTGGGACTATCAG 1380  
Db 1321 TACTTCTTTTCAACAAACCCACTAAATCTATCTCCAGAGTATTGCTGGGACTATCAG 1380  
Qy 1381 ATAGGCTGCTTCAGATATTAAAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAAT 1440  
Db 1381 ATAGGCTGCTTCAGATATTAAAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAAT 1440  
Qy 1441 TTGGTTAGAAATTAATCTGACTTCAAAATTTGATGGAACCTTGACACTATTCTTAACCA 1500  
Db 1441 TTGGTTAGAAATTAATCTGACTTCAAAATTTGATGGAACCTTGACACTATTCTTAACCA 1500

RESULT 2  
US-08-823-489-3  
; Sequence 3, Application US/08823489  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUBSTADT,  
; ADDRESSER: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/823,489  
; FILING DATE: 25-MAR-1997  
; CLASSIFICATION: 435  
; PRIORITY DATA:  
; PRIORITY NUMBER: 07/914,281

; FILING DATE: 20-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1500 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cdna  
; ANTI-SENSE: NO  
; US-08-823-489-3

Query Match 100.0%; Score 1500; DB 12; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCCCTTTGAGACTCTTCTTGGAAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Db 1 CCTTCCCTTTGAGACTCTTCTTGGAAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Qy 61 CTCAGGCTCTGGAGAGTTGGAAACCTGTACCTTCTCTCTGCTGAGCCCTGCTTCTT 120  
Db 61 CTCAGGCTCTGGAGAGTTGGAAACCTGTACCTTCTCTCTGCTGAGCCCTGCTTCTT 120  
Qy 121 AGCAGGCCAGAGCTCGACAGAACTCGGTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTT 180  
Db 121 AGCAGGCCAGAGCTCGACAGAACTCGGTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTT 180  
Qy 181 CTGACGATGAGGCTGACTTTTGAAGTCAAGAGATCTGCTTACCCAGTCTCCTTGAATTA 240  
Db 181 CTGACGATGAGGCTGACTTTTGAAGTCAAGAGATCTGCTTACCCAGTCTCCTTGAATTA 240  
Qy 241 AGCCCTGTACTACATTTGCTGACCTAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Db 241 AGCCCTGTACTACATTTGCTGACCTAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Qy 301 CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAAACAAGTCTTCCATCTCTCAAGA 360  
Db 301 CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAAACAAGTCTTCCATCTCTCAAGA 360  
Qy 361 TCTGGATCAGAGAGAAATAATGAATGTCAAGGAAAGTAAATCTCTGTTGATGCTGATTT 420  
Db 361 TCTGGATCAGAGAGAAATAATGAATGTCAAGGAAAGTAAATCTCTGTTGATGCTGATTT 420  
Qy 421 GTCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
Db 421 GTCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
Qy 481 AACAGATGGCAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
Db 481 AACAGATGGCAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
Qy 541 CAAGAAGACAACGTAG 600  
Db 541 CAAGAAGACAACGTAG 600  
Qy 601 CCTCAGCTATGGGACTTGGTTCATCCAAAGACCGCCCGGATGTTTTCAGACTGACCCCG 660  
Db 601 CCTCAGCTATGGGACTTGGTTCATCCAAAGACCGCCCGGATGTTTTCAGACTGACCCCG 660  
Qy 661 TGGAGGCGCGGATTTGTTGGGAGAGGACCTTATGACACAGCTCTGCTGGAGAGAGTACTAC 720  
Db 661 TGGAGGCGCGGATTTGTTGGGAGAGGACCTTATGACACAGCTCTGCTGGAGAGAGTACTAC 720  
Qy 721 GCCACACAGAAACTCACTGTTGGGCTGACAGTGTGTTGCTCTGGGAAAGTACATTTGAGCAT 780

Db 721 GCCACACAGAAATCACTGTGGGCTGACAGTGTTCCTGTGGAAAGTACATTGACCAT 780  
QY 781 TACTTAGAAGACTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTCATA 840  
Db 781 TACTTAGAAGACTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTCATA 840  
QY 841 TTTTACGTATGATAGACGACACTCCCGGATCCCTGCTGCTGACCTGAACCCCTCTACAT 900  
Db 841 TTTTACGTATGATAGACGACACTCCCGGATCCCTGCTGCTGACCTGAACCCCTCTACAT 900  
QY 901 TCTTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGGTGGGAGATATCAGCATGATCGC 960  
Db 901 TCTTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGGTGGGAGATATCAGCATGATCGC 960  
QY 961 ATCAAGACCAATTTGGGAGACATCTTGGCCACATCCAGCACGAGGTGCGATTCCTCTTC 1020  
Db 961 ATCAAGACCAATTTGGGAGACATCTTGGCCACATCCAGCACGAGGTGCGATTCCTCTTC 1020  
QY 1021 TGATGACGTGGATCAAGTCTTTCAAGACAACCTTGGGGGTGAAACCTCTGGGCCAGCTG 1080  
Db 1021 TGATGACGTGGATCAAGTCTTTCAAGACAACCTTGGGGGTGAAACCTCTGGGCCAGCTG 1080  
QY 1081 GTAGCAGCTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
Db 1081 GTAGCAGCTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
QY 1141 CGGGAATCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
Db 1141 CGGGAATCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
QY 1201 ATTTTGGAGAACGCTACTACATCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCT 1260  
Db 1201 ATTTTGGAGAACGCTACTACATCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCT 1260  
QY 1261 CTCAGSACAAGAACATGACATGACAGCCAGTGGCATGATGAGAGCCACCTCAACAAA 1320  
Db 1261 CTCAGSACAAGAACATGACATGACAGCCAGTGGCATGATGAGAGCCACCTCAACAAA 1320  
QY 1321 TACTTCTCTTTTCAACAAACCCACTAAATCTATCTCCAGAGTATTTGCTGGGACTATCAG 1380  
Db 1321 TACTTCTCTTTTCAACAAACCCACTAAATCTATCTCCAGAGTATTTGCTGGGACTATCAG 1380  
QY 1381 ATAGGCTGCTTTCAGATATTTAAAGTGTCAAGGTAGCTTGGCAGACAAAAGATATAAT 1440  
Db 1381 ATAGGCTGCTTTCAGATATTTAAAGTGTCAAGGTAGCTTGGCAGACAAAAGATATAAT 1440  
QY 1441 TTGGTTAGAAATATGCTGCTCACTTCAAAATTTGATGGAACCTTGACACTATTTCTAACCA 1500  
Db 1441 TTGGTTAGAAATATGCTGCTCACTTCAAAATTTGATGGAACCTTGACACTATTTCTAACCA 1500

RESULT 3

US-09-593-316-9  
; Sequence 9, Application US/09593316  
; GENERAL INFORMATION:  
; APPLICANT: Clark Mr., John  
; APPLICANT: Denning, Chris  
; TITLE OF INVENTION: Animal Tissue For Xenotransplantation  
; FILE REFERENCE: 730/002  
; CURRENT APPLICATION NUMBER: US/09/593,316  
; PRIOR FILING DATE: 2000-06-13  
; PRIOR FILING DATE: 2000-05-15  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 1500  
; TYPE: DNA  
; ORGANISM: Musca sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (277)..(1461)

US-09-593-316-9  
Query Match 100.0%; Score 1500; DB 22; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTTCCCTTGTAGACTCTCTTCTGGAATGAGAAGTACCGATTCTGCTGCAAGACCTCGCGCT 60  
Db 1 ccttccctctgtagactctctcttggaatgagaagtacagattctgctgaagacctcgagct 60  
QY 61 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCTCTTTCCTCTGCTGAGCCCTGCTCTCTT 120  
Db 61 ctcaggctctgggagttggaaacctgtacctctctctctctctctctctctctctctctct 120  
QY 121 AGGAGGCGGACAGAGCTGACAGAACTCGGTTGCTTGTGTTTGTGTTTGTGAGGGGAACACAG 180  
Db 121 aggaggcgagagctgacagaaactcggcttgccttgccttgccttgccttgccttgccttgc 180  
QY 181 CTGACGATGAGGCTGACTTTTGAACCTCAAGAGATCTGCTTACCCGAGTCTCTTGGAAATTA 240  
Db 181 ctgacgatgaggctgactttgaaactcaagagatctgcttaccctcagctctctctggaattaa 240  
QY 241 AGGCTGTACTACATTTGCTTGGACCTTAAGATTTTCATGATCAGTATGCTTCAAGATCTC 300  
Db 241 aggcctgtactacatttgcctggacctaaatttctcatgatacactatgcttcaagatctc 300  
QY 301 CATGTCAACAGATCTCCATGTCAAGATCCAAGTCAGAAACAAGTCTTCCATCTCTCAAGA 360  
Db 301 catgtcaacaagatctccatgtccaagtccaagtcagaacaagaagcttccctctctctcaaga 360  
QY 361 TCTGGATCACAGGAGAAATAATGAATGTCAAGGAAAGTAATCTCTGTTGATGCTCTGAT 420  
Db 361 tctgga tcacagagagaaaaataatgaatgtccaagggaaaagtaacccctgtgtgctgctgatt 420  
QY 421 GTCTCAACCGTGTGCTGCTGCTTGGGAATATGTCACAGAAATTCACAGAGTTGGTGAG 480  
Db 421 gtctcaaccgtgtgtgctgcttgggaatatgtccaagggaaaagtaacccctgtgtgctgctgatt 480  
QY 481 AACAGATGGGAGAGGACTGTGCTGCCAAGCTGGTTTTAAATAATGGGACCCACAGATTAT 540  
Db 481 aacaga tggcagaagagactggtgttcccaagctggttttaaaaaatgggacccacagttat 540  
QY 541 CAAGAAGACAACTGATAGAGGACGAGAGAAAGGTTAGAAATGGAGATCGGATTTGAAGAG 600  
Db 541 caagaagacaactgataagagacgagagagaaaggttagaaatggagatcgcatttgaagag 600  
QY 601 CCTCAGCTATGGGACTGTTCAATCCAAAGAACCCCGGATGTTTTCACAGTGCACCCCG 660  
Db 601 cctcagcta tgggactggttcaa tccaaagaaccccgccggatgttttgcagagtgaccctcg 660  
QY 661 TGAAGGCGCGGATTTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAGTACTATAC 720  
Db 661 tggaaagcgcgattgttgggaagagcatttatgacacagctctgctggaaaagtactac 720  
QY 721 GCCACACAGAACTCACTGTGGGCTGACAGTGTGTTGCTGTGGGAAAGTACATTTGAGCAT 780  
Db 721 gccacacagaactcactgtgggctgacagttgtgctgtgggaaagatcatattgagcat 780  
QY 781 TACTTAGAAGACTTCTGAGGCTGCTGACATGTTTTCATGTTGGCCATCGGGTCATA 840  
Db 781 tacttagaagactctctggagctgctgacatgtacattca tgg tgggcca tgggggtcata 840  
QY 841 TTTTACGTATGATAGACGACACTCCCGGATCCCTGCTGCTGCAACCTCTACAT 900  
Db 841 ttttacgtatgatagacgacacctcccgatgctgctgacacctgaacctctacat 900  
QY 901 TCTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGGTGGCAGGATATCAGCATGATCGGC 960  
Db 901 tccttaaaagctcttggatgatacaggtctgagaagaggtggcaggatatacagcatgagctgc 960  
QY 961 ATGAAGACCAATTTGGGAGGACATCTTGGCCACATCCAGCACGAGGTGCGACTTCTCTTC 1020  
Db 961 atgaagaccaatttgggagggacatcttggccacatccagcacagggctgacttctctcttc 1020

Db 961 atgaagaccattgaggagacacatcctgccccacatccacagcagaggtcgacttcctctc 1020  
Qy 1021 TCATCGAGCGTGGATCAAGTCCTTTCAAGACAACTTCGGGGTGGAACTCTGGCCGACGTG 1080  
Db 1021 tgcattgagctgagatcaagctcttcaagacaacttcggggtggaactctggccagctg 1080  
Qy 1081 GTAGCACAGCTCCAGCGCTGGGTGGGTACAGGCCAGTCCCGGAGAGTTACCTATGAGAG 1140  
Db 1081 gtagcacagctccaggcctgg tggLacaaggccagtcccagagaagttcacctatgagag 1140  
Qy 1141 CGGGAACCTCGCGCGCTGACATTCATTCGAGAGGGGATTTTACTACCACCGCGCC 1200  
Db 1141 cgggaacctcgcgcgcgcacattccattcggagaggggattttactaccacgcgcgc 1200  
Qy 1201 ATTTTGGAGAACCGCTACTCACATTTCAACCTCACCGAGGTGCTTTAAGGGGATC 1260  
Db 1201 attttggaggaacgcctactcacattctcaacctcaccagagtgctttaaggggatc 1260  
Qy 1261 CTCACGACGAAGAATGACATAGAGCCCGAGTGGCATGTGAGAGCCACCTCAACAAA 1320  
Db 1261 ctccaggacaagaacatgacatagaagcccgatggcatgatgagagccacctcaacaaa 1320  
Qy 1321 TACTTCCCTTTCAACAAACCCACTAAATCCCTATCTCCAGACTATTGCTGGGCACTATCAG 1380  
Db 1321 tacttccctttcaacaacccccactaaaatccctatctccagagtatgtgctgggactatcag 1380  
Qy 1381 ATAGGCCTGCCCTCAGATATTAAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAAT 1440  
Db 1381 ataggcctgcctcagatattaaaagtgcaagtgatgcttggcagacaaaagagataat 1440  
Qy 1441 TTGGTTAGAAATAATGCTGACTTCAAATTTGATGGAAACTTGACACTATTTTCTAAACA 1500  
Db 1441 ttggttagaaaataatgctgacttcaaatgtgagaaacttgacactatttctlaacca 1500

RESULT 4  
US-09-863-475-3  
; Sequence 3, Application US/09863475  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/863,475  
; FILING DATE: 24-May-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/914,281  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1500 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; ANTI-SENSE: NO  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-863-475-3

Query Match	Best Local Similarity	Score	DB	Length	Mismatches	Indels	Gaps
Qy 1	1	100.0%	1500;	1500;	0;	0;	0;
Db 1	1	100.0%	1500;	1500;	0;	0;	0;
Qy 61	61	100.0%	1500;	1500;	0;	0;	0;
Db 61	61	100.0%	1500;	1500;	0;	0;	0;
Qy 121	121	100.0%	1500;	1500;	0;	0;	0;
Db 121	121	100.0%	1500;	1500;	0;	0;	0;
Qy 181	181	100.0%	1500;	1500;	0;	0;	0;
Db 181	181	100.0%	1500;	1500;	0;	0;	0;
Qy 241	241	100.0%	1500;	1500;	0;	0;	0;
Db 241	241	100.0%	1500;	1500;	0;	0;	0;
Qy 301	301	100.0%	1500;	1500;	0;	0;	0;
Db 301	301	100.0%	1500;	1500;	0;	0;	0;
Qy 361	361	100.0%	1500;	1500;	0;	0;	0;
Db 361	361	100.0%	1500;	1500;	0;	0;	0;
Qy 421	421	100.0%	1500;	1500;	0;	0;	0;
Db 421	421	100.0%	1500;	1500;	0;	0;	0;
Qy 481	481	100.0%	1500;	1500;	0;	0;	0;
Db 481	481	100.0%	1500;	1500;	0;	0;	0;
Qy 541	541	100.0%	1500;	1500;	0;	0;	0;
Db 541	541	100.0%	1500;	1500;	0;	0;	0;
Qy 601	601	100.0%	1500;	1500;	0;	0;	0;
Db 601	601	100.0%	1500;	1500;	0;	0;	0;
Qy 661	661	100.0%	1500;	1500;	0;	0;	0;
Db 661	661	100.0%	1500;	1500;	0;	0;	0;
Qy 721	721	100.0%	1500;	1500;	0;	0;	0;
Db 721	721	100.0%	1500;	1500;	0;	0;	0;
Qy 781	781	100.0%	1500;	1500;	0;	0;	0;
Db 781	781	100.0%	1500;	1500;	0;	0;	0;
Qy 841	841	100.0%	1500;	1500;	0;	0;	0;



QY 721 GCCACACAGAACTACTGTGGGCTGACAGTGTTCGTGTGGAAAGTACATTTGAGCAT 780  
Db 721 GCCACACAGAACTACTGTGGGCTGACAGTGTTCGTGTGGAAAGTACATTTGAGCAT 780  
QY 781 TACTTTAGAAGACTTTCTGGAGTCTGTGACATCTACTTCTATGTTGGCCATCGGTCATA 840  
Db 781 TACTTTAGAAGACTTTCTGGAGTCTGTGACATCTACTTCTATGTTGGCCATCGGTCATA 840  
QY 841 TTTTACGTCATGATAGACACACCTCCCGGATGCCTGTCTGTCACCTTGAACCCCTCTACAT 900  
Db 841 TTTTACGTCATGATAGACACACCTCCCGGATGCCTGTCTGTCACCTTGAACCCCTCTACAT 900  
QY 901 TCTTTACAAGCTTTGAGATCAGGTCTGAGAAGAGGTGGCAGGATATCAGCATGATGCCG 960  
Db 901 TCTTTACAAGCTTTGAGATCAGGTCTGAGAAGAGGTGGCAGGATATCAGCATGATGCCG 960  
QY 961 ATCAAGACCAATTCGGGAGCAGCATCTGGCCACATCCAGCAGGTCGACTTCTCTTC 1020  
Db 961 ATCAAGACCAATTCGGGAGCAGCATCTGGCCACATCCAGCAGGTCGACTTCTCTTC 1020  
QY 1021 TGCATGGACGTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAACTCTTGGGCCAGCTG 1080  
Db 1021 TGCATGGACGTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAACTCTTGGGCCAGCTG 1080  
QY 1081 GTAGCAGCTCAGGCTGGTGTGTAAGGCCAGTCCCGAGAGTTCACCTATGAGAGG 1140  
Db 1081 GTAGCAGCTCAGGCTGGTGTGTAAGGCCAGTCCCGAGAGTTCACCTATGAGAGG 1140  
QY 1141 CGGGAAGTCTCGGCCGCTACATTCCTCCATTCGGGAGGGGATTTTACTACACGCGGCC 1200  
Db 1141 CGGGAAGTCTCGGCCGCTACATTCCTCCATTCGGGAGGGGATTTTACTACACGCGGCC 1200  
QY 1201 ATTTTGGAGAACGCTACTACATCTTCAACCTTCACAGGAGTGTCTTAAGGGGATC 1260  
Db 1201 ATTTTGGAGAACGCTACTACATCTTCAACCTTCACAGGAGTGTCTTAAGGGGATC 1260  
QY 1261 CTCAGGACAAGAACATGACATAGAGCCAGTGGCATGATGAGAGCCACCTCAACAAA 1320  
Db 1261 CTCAGGACAAGAACATGACATAGAGCCAGTGGCATGATGAGAGCCACCTCAACAAA 1320  
QY 1321 TACTTCTTTCAACAAACCCACTAAATCTATCTCCAGAGTATTGCTGGGACTATCAG 1380  
Db 1321 TACTTCTTTCAACAAACCCACTAAATCTATCTCCAGAGTATTGCTGGGACTATCAG 1380  
QY 1381 ATAGGCTGCTTTCAGATATTAAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAAT 1440  
Db 1381 ATAGGCTGCTTTCAGATATTAAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAAT 1440  
QY 1441 TTGGTTAGAAATAATGCTGACTTCAAAATTTGTGATGGAACCTTGACACTATTTCTAACCA 1500  
Db 1441 TTGGTTAGAAATAATGCTGACTTCAAAATTTGTGATGGAACCTTGACACTATTTCTAACCA 1500

RESULT 6

US-09-946-034-1

GENERAL INFORMATION:

APPLICANT: Cooper, David K.C.

Koren, Eugen

TITLE OF INVENTION: GENETICALLY ENGINEERED ANIMALS FOR USE

AS ORGAN DONORS

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst

STREET: 1100 Peachtree Street, Suite 2800

CITY: Atlanta

STATE: GA

COUNTRY: USA

ZIP: 30309-4530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/946,034  
FILING DATE: 04-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,817  
FILING DATE: 20-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: BMC100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)-815-6508  
TELEFAX: (404)-815-6555  
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
VOLUME: 86  
ISSUE: November  
PAGES: 8227-8231  
DATE: 1989  
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 1500  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-946-034-1

Query Match 100.0%; Score 1500; DB 35; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTTCCCTTTAGACTCTCTTGGAAATGAGAACTACCGATTCGCTCAAGACCTCGCGCT 60  
Db 1 CTTTCCCTTTAGACTCTCTTGGAAATGAGAACTACCGATTCGCTCAAGACCTCGCGCT 60  
QY 61 CTCAGGCTCTGGAGATTTGGAACTGTACCTTCTCTTCTCTCTGCTGAGCCCTGCTCTCT 120  
Db 61 CTCAGGCTCTGGAGATTTGGAACTGTACCTTCTCTTCTCTCTGCTGAGCCCTGCTCTCT 120  
QY 121 AGGAGCCAGAGCTCGACAGAACTCGGTTGCTGTTTGTGTTTGGAGGGAACACAG 180  
Db 121 AGGAGCCAGAGCTCGACAGAACTCGGTTGCTGTTTGTGTTTGGAGGGAACACAG 180  
QY 181 CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCTCGAATTA 240  
Db 181 CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCTCGAATTA 240  
QY 241 AGGCTGTACTACATTTGCTGGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Db 241 AGGCTGTACTACATTTGCTGGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
QY 301 CATGTCAACAGATCTCCATGTCAAGATCCAAAGTCAAGAACAAAGTCTTCCATCTCAAGA 360  
Db 301 CATGTCAACAGATCTCCATGTCAAGATCCAAAGTCAAGAACAAAGTCTTCCATCTCAAGA 360  
QY 361 TCTGGATCAGAGGAGAAATAATCAATGTCAAGGGAAGTAATCTCTGTCATGCTGATT 420  
Db 361 TCTGGATCAGAGGAGAAATAATCAATGTCAAGGGAAGTAATCTCTGTCATGCTGATT 420  
QY 421 GTCTCAACCGTGTGCTGCTTTGGGAATATGTCAACAGAAATCCAGAGGTTGGTCAG 480  
Db 421 GTCTCAACCGTGTGCTGCTTTGGGAATATGTCAACAGAAATCCAGAGGTTGGTCAG 480  
QY 481 AACAGATGGCAGAGGAGTGGTGTGCTGCTTCCCAAGCTGGTTTAAATGGGACCCACAGTTAT 540  
Db 481 AACAGATGGCAGAGGAGTGGTGTGCTGCTTCCCAAGCTGGTTTAAATGGGACCCACAGTTAT 540  
QY 541 CAAGACACAACTGTAGAGGACGAGAGAAAGGTTAGAAATGAGATCGGATTTGAAGAG 600  
Db 541 CAAGACACAACTGTAGAGGACGAGAGAAAGGTTAGAAATGAGATCGGATTTGAAGAG 600  
QY 601 CCTCAGCTATGGGACTGTTCAATCCAAAGAACCCCGGATGTTTTCACAGTCAACCCCG 660  
Db 601 CCTCAGCTATGGGACTGTTCAATCCAAAGAACCCCGGATGTTTTCACAGTCAACCCCG 660  
QY 661 TGGAGGCGCGGATGTTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAGTACTTAC 720





```

1 INFORMATION FOR SEQ ID NO: 9:
2 SEQUENCE CHARACTERISTICS:
3     LENGTH: 3450
4     TYPE: nucleic acid
5     STRANDEDNESS: double strand
6     TOPOLOGY: linear
7     MOLECULE TYPE: cDNA
8     DESCRIPTION:
9     HYPOTHETICAL: no
10    ANTI-SENSE: no
11    FRAGMENT TYPE:
12    ORIGINAL SOURCE:
13        ORGANISM: Mus musculus
14        STRAIN:
15    INDIVIDUAL ISOLATE:
16    IMMEDIATE SOURCE:
17    LIBRARY:
18    CLONE:
19    POSITION IN GENOME:
20    CHROMOSOME/SEGMENT:
21    FEATURE:
22        NAME/KEY: Nucleotide sequence
23    PUBLICATION INFORMATION:
24        AUTHORS: Joctasse, D. H.
25        AUTHORS: Shaper, N. L.
26        AUTHORS: Kim, D.
27        AUTHORS: Van den Eljnden,
28        AUTHORS: Shaper, J. H.
29        TITLE: Murine alpha-1,3-galactosyl
30        JOURNAL: The Journal of Biological Chemistry
31        VOLUME: 267
32        ISSUE: 8
33        PAGES: 5534-5541
34        DATE: 1992
35 US-08-188-607B-9

```

RESOL 9  
IIS-08-188-607B-9

US-08-188-607B-9  
Sequence 9, Application US/08188607B  
GENERAL INFORMATION:  
APPLICANT: d'Apice, Anthony J.F.  
APPLICANT: Pearce, Martin J.  
APPLICANT: Robins, Allan J.  
APPLICANT: Crawford, Robert J.  
APPLICANT: Rathjen, Peter D.  
TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF HYPERACUTE  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patterson & Keough, P.A.  
STREET: 527 Marquette Avenue South  
CITY: Minneapolis  
STATE: Minnesota  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh System 7.0  
SOFTWARE: Macintosh Text File  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,607B  
FILING DATE: 27-01-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark S. Ellinger  
REGISTRATION NUMBER: 34,812  
REFERENCE/DOCKET NUMBER: 1175.04-US-01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612/349-5740  
TELEFAX: 612/349-9266

```
Qy 462 -----AATTCAGAGGTGGTGAGAACAGATGGCAGAGGACTGGTG 503
Db 543 GTGGATATATACACAAAAATTCAGAGGTGGTGAGAACAGATGGCAGAGGACTGGTG 602
Qy 504 GTTCCCAAGCTGGTTTAAAAATGGACCCACAGTTATCAAGAACACACGCTAGAGGACG 563
Db 603 GTTCCCAAGCTGGTTTAAAAATGGACCCACAGTTATCAAGAACACACGCTAGAGGACG 662
Qy 564 GAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAA 623
Db 663 GAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAA 722
Qy 624 TCCAAAGAACGCCCGGATGTTTGGACAGTGACCCCTGGAAAGCGCGGATGTTGGGGA 683
Db 723 TCCAAAGAACGCCCGGATGTTTGGACAGTGACCCCTGGAAAGCGCGGATGTTGGGGA 782
Qy 684 AGGCACCTTATGACACAGCTCTGCTGGAAAGTACTACGCCACACAGAACTCACTGTGGG 743
Db 783 AGGCACCTTATGACACAGCTCTGCTGGAAAGTACTACGCCACACAGAACTCACTGTGGG 842
Qy 744 GCTGACAGTGTGCTGTGGGAAAGTACATTTAGCATTACTTTAGAAGACTTTCTGGAGTC 803
Db 843 GCTGACAGTGTGCTGTGGGAAAGTACATTTAGCATTACTTTAGAAGACTTTCTGGAGTC 902
Qy 804 TGCTGACATGACTTTCATGTTGGCCATCGGGTGCATATTTTACGTCATGATAGACGAC 863
Db 903 TGCTGACATGACTTTCATGTTGGCCATCGGGTGCATATTTTACGTCATGATAGACGAC 962
Qy 864 CTCCCGGATGCTGCTGACAGCTGAACCTCTACATTCCTTACAAGTCTTTGAGATCAG 923
Db 963 CTCCCGGATGCTGCTGACAGCTGAACCTCTACATTCCTTACAAGTCTTTGAGATCAG 1022
Qy 924 GTCTGAGAAGAGTGGCAGGATATCAGCATGATCGCATGAAGACCACTTTGGGAGCACAT 983
Db 1023 GTCTGAGAAGAGTGGCAGGATATCAGCATGATCGCATGAAGACCACTTTGGGAGCACAT 1082
Qy 984 CTTGGCCCATCATCAGACAGAGTGCAGTCTCTCTTCTGATGAGACGTGGATCAAGTCTT 1043
Db 1083 CTTGGCCCATCATCAGACAGAGTGCAGTCTCTCTTCTGATGAGACGTGGATCAAGTCTT 1142
Qy 1044 TCAAGACAACTTCGGGGTGAAACTCTGGCCAGCTGGTAGCAGAGCTCCAGGCGCTGGTG 1103
Db 1143 TCAAGACAACTTCGGGGTGAAACTCTGGCCAGCTGGTAGCAGAGCTCCAGGCGCTGGTG 1202
Qy 1104 GTAAAGGCCAGTCCCAGAGATTTACCTATGAGAGCGGGAAGTCTGCGGCCGCTACAT 1163
Db 1203 GTAAAGGCCAGTCCCAGAGATTTACCTATGAGAGCGGGAAGTCTGCGGCCGCTACAT 1262
Qy 1164 TCCATTCCGAGAGGGGATTTTACTACCGCGGCCATTTTGGAGGAAGCGCTACTCA 1223
Db 1263 TCCATTCCGAGAGGGGATTTTACTACCGCGGCCATTTTGGAGGAAGCGCTACTCA 1322
Qy 1224 CATTTCAACCTCACAGGAGTCTTTTAAAGGGATCTCCAGGACAAAGAACATGACAT 1283
Db 1323 CATTTCAACCTCACAGGAGTCTTTTAAAGGGATCTCCAGGACAAAGAACATGACAT 1382
Qy 1284 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCAC 1343
Db 1383 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCAC 1442
Qy 1344 TAAATCCTATCTCCAGATATTGCTGGGACTATCAGATAGGCGCTGCTTCAGATATTA 1403
Db 1443 TAAATCCTATCTCCAGATATTGCTGGGACTATCAGATAGGCGCTGCTTCAGATATTA 1502
Qy 1404 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAATTTGGTTAGAAATATGTCGACT 1463
Db 1503 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAATTTGGTTAGAAATATGTCGACT 1562
Qy 1464 TCAATTTGTGATGAAACTTTGACACTATTTCT 1495
Db 1563 TCAATTTGTGATGAAACTTTGACACTATTTCT 1594
```

```
RESULT 9
US-08-378-617-9
: Sequence 9, Application US/08378617
: GENERAL INFORMATION:
: APPLICANT: d'Apice, Anthony J.F.
: APPLICANT: Pearse, Martin J.
: APPLICANT: Robins, Allan J.
: APPLICANT: Crawford, Robert J.
: APPLICANT: Rathjen, Peter D.
: TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF
: TITLE OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 120 South Sixth Street, Suite 2500
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30B
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/378,617
: FILING DATE: 26-JAN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Ellinger, Mark S.
: REGISTRATION NUMBER: 34,812
: REFERENCE/DOCKET NUMBER: 06868/005001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (612) 335-5070
: TELEFAX: (612) 288-9696
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3450 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
US-08-378-617-9
```

```
Query Match 93.9% Score 1408.4; DB 7; Length 3450;
Best Local Similarity 96.8% Pred. No. 0;
Matches 1483; Conservative 0; Mismatches 11; Indels 38; Gaps 3;
```

```
Qy 1 CCTTCCCTTTAGACTCTTCTTGGAAATGAGAAATACCGATTCTGCTGAAGACCTCGCGCT 60
Db 64 CCTTCCCTTTAGACTCTTCTTGGAAATGAGAAATACCGATTCTGCTGAAGACCTCGCGCT 123
Qy 61 CTCAGGCTCTGGAGTTGGAACCTT-GTACCTTCTCTTCTCTCTGCTGAGCCCTGCTCTCT 119
Db 124 CTCAGGCTCTGGAGTTGGAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 183
Qy 120 TAGCAGGCCAGAGCTCGACAGAACTCGGTTGCTTTGCTTTGCTTTGGAGGGAACACA 179
Db 184 TCGCAGGCCAGAGCTCGACAGAACTCGGTTGCTTTGCTTTGCTTTGGAGGGAACACA 243
Qy 180 GCTGACGATGAGGCTGACTTTTGAAGTCAAGAGATCTGCTTACCCAGTCTCTCTGGAATTA 239
Db 244 GCTGACGATGAGGCTGACTTTTGAAGTCAAGAGATCTGCTTACCCAGTCTCTCTGGAATTA 303
Qy 240 AAGCCCTGCTACTACTATTGCTCGCCGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCT 299
Db 304 AAGCCCTGCTACTACTATTGCTCGCCGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCT 362
Qy 300 CCATGTCAACAAAGATCTCCATGTCAAGATCCAAGTCAAGAAACAAAGTCTTCCATCTCAAG 359
Db 363 CCATGTCAACAAAGATCTCCATGTCAAGATCCAAGTCAAGAAACAAAGTCTTCCATCTCAAG 422
```

QY 360 ATCTGGATCACAGGAGAAATAATGAATGTCAAGGAAAGTAATCTCTTGTATGCTGAT 419  
Db 423 ATCTGGATCACAGGAGAAATAATGAATGTCAAGGAAAGTAATCTCTTGTATGCTGAT 482  
QY 420 TCTCTCAACCGTGTGTCTGTCTTGGGAATATGTCAACAG----- 461  
Db 483 TGTCTCAACCGTGTGTCTGTCTTGGGAATATGTCAACAGCCCAAGACGGCTCTTCTTT 542  
QY 462 -----AATTCAGAGGTGTGTGAGAACAGATGGCAGAAAGACTGTGTG 503  
Db 543 GTGGATATATACACAAAATTCAGAGGTGTGTGAGAACAGATGGCAGAAAGACTGTGTG 602  
QY 504 GTTCCCAAGCTGTTTAAATATGGACCCACACAGTTATCAAGAGACAACTAGAGGAGC 563  
Db 603 GTTCCCAAGCTGTTTAAATATGGACCCACACAGTTATCAAGAGACAACTAGAGGAGC 562  
QY 564 GAGAGAAAGGTGAGAAATGGAGATCGCATTCGAAGAGCTCAGCTATGGAGCTGTTCAA 623  
Db 663 GAGAGAAAGGTGAGAAATGGAGATCGCATTCGAAGAGCTCAGCTATGGAGCTGTTCAA 722  
QY 624 TCCAAGAAACCCCGGATGTTTTCACAGTGACCCCGTGGAAAGCGCCGATTTCTGTGGGA 683  
Db 723 TCCAAGAAACCCCGGATGTTTTCACAGTGACCCCGTGGAAAGCGCCGATTTCTGTGGGA 782  
QY 684 AGGCACCTTATGACACAGCTCTCTGTGAAAGTACTACGCCACACAGAAACTCACTGTGG 743  
Db 783 AGGCACCTTATGACACAGCTCTCTGTGAAAGTACTACGCCACACAGAAACTCACTGTGG 842  
QY 744 GCTGACAGTGTTCGTGCGGAAGTACATTCGACATTAATAGAGACTTTCTGGAGTC 803  
Db 843 GCTGACAGTGTTCGTGCGGAAGTACATTCGACATTAATAGAGACTTTCTGGAGTC 902  
QY 804 TGCTGACATGTACTTCATCGTTGGCCATCGGGTCATATTTTACGTCATGATAGACACAC 863  
Db 903 TGCTGACATGTACTTCATCGTTGGCCATCGGGTCATATTTTACGTCATGATAGACAC 962  
QY 864 CTCCCGGATGCCTGTGCTGCACCTGAAACCTCTACATTCCTTACAGTCTTTGAGATCAG 923  
Db 963 CTCCCGGATGCCTGTGCTGCACCTGAAACCTCTACATTCCTTACAGTCTTTGAGATCAG 1022  
QY 924 GTCGAGAGAGGTGCGCAGATATCAGCATGATGCGCATGAGACCATTTGGGAGCACAT 983  
Db 1023 GTCGAGAGAGGTGCGCAGATATCAGCATGATGCGCATGAGACCATTTGGGAGCACAT 1082  
QY 984 CTGGCCCCACATCCAGCAGAGTTCGACTTCCTTCTTCGATGGAGCTGGATCAAGTCTT 1043  
Db 1083 CTGGCCCCACATCCAGCAGAGTTCGACTTCCTTCTTCGATGGAGCTGGATCAAGTCTT 1142  
QY 1044 TCAAGACAATTCGGGGTGAAACTCTGGGCCAGCTGGTGTAGACAGCTCCAGGCCCTGGTG 1103  
Db 1143 TCAAGACAATTCGGGGTGAAACTCTGGGCCAGCTGGTGTAGACAGCTCCAGGCCCTGGTG 1202  
QY 1104 GTACAGGCCAGTCCCGAGAAGTTACCTATGAGAGCGGGAACCTCTCGGCCGCTACAT 1163  
Db 1203 GTACAGGCCAGTCCCGAGAAGTTACCTATGAGAGCGGGAACCTCTCGGCCGCTACAT 1262  
QY 1164 TCATTTCGAGAGGGGATTTTACTACACCGCGGCATTTTGGAGAACGCTTACTCA 1223  
Db 1263 TCATTTCGAGAGGGGATTTTACTACACCGCGGCATTTTGGAGAACGCTTACTCA 1322  
QY 1224 CATTTCTCAACCTCACAGGAGTGTCTTTAAGGGATCTCCAGGACGAAGAAATGACAT 1283  
Db 1323 CATTTCTCAACCTCACAGGAGTGTCTTTAAGGGATCTCCAGGACGAAGAAATGACAT 1382  
QY 1284 AGAAGCCAGTGGCATGATGAGAGCCACCTTCAACAATACTCTCTTTTCAACAACCCAC 1343  
Db 1383 AGAAGCCAGTGGCATGATGAGAGCCACCTTCAACAATACTCTCTTTTCAACAACCCAC 1442  
QY 1344 TAAATTCCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCCCTTCAGATATTA 1403  
Db 1443 TAAATTCCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCCCTTCAGATATTA 1502  
QY 1404 AAGTGTCAAAGGTAGCTTGGCAGACAAAAGAGTATAATTTGGTTAGAAAATAATCTCGACT 1463

Db 1503 AAGTGTCAAAGGTAGCTTGGCAGACAAAAGAGTATAATTTGGTTAGAAAATAATCTCGACT 1562  
QY 1464 TCAAAATTTGATGGAAACTTGACACTATTCT 1495  
Db 1563 TCAAAATTTGATGGAAACTTGACACTATTACT 1594  
RESULT 10  
US-08-984-900-9  
; Sequence 9, Application US/08984900  
; GENERAL INFORMATION:  
; APPLICANT: d'Apice, Anthony J.F.  
; APPLICANT: Pearce, Martin J.  
; APPLICANT: Robins, Allan J.  
; APPLICANT: Crawford, Robert J.  
; APPLICANT: Rathjen, Peter D.  
; TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF  
; TITLE OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 120 South Sixth Street, Suite 2500  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/984,900  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/378,617  
; FILING DATE: 26-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 06868/005001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (612) 335-5070  
; TELEFAX: (612) 288-9696  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3450 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-984-900-9

Query Match 93.9%; Score 1408.4; DB 13; Length 3450;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 1483; Conservative 0; Mismatches 11; Indels 38; Gaps 3;  
QY 1 CTTTCCCTCTGTAGACTCTTCTTGAATGAGAAGTACCGATTTCTGCTGAAGACCTCGCGCT 60  
Db 64 CTTTCCCTCTGTAGACTCTTCTTGAATGAGAAGTACCGATTTCTGCTGAAGACCTCGCGCT 123  
QY 61 CTCAGGCTCTGGAGTGGAAACCT-CTACCTTCCCTTCTCTCTCTCTCTCTCTCTCTCTCT 119  
Db 124 CTCAGGCTCTGGAGTGGAAACCTCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 183  
QY 120 TAGCGAGCCAGAGCTCGACAGAACTCGGTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 179  
Db 184 TCGGCAGCCAGAGCTCGACAGAACTCGGTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 243  
QY 180 GCTGACGATGAGGCTGACTTTTGAACCTCAAGAGATCTGCTTACCCCACTCTCTCTCTCTCT 239

Db 244 GCTGAGGATGAGGCTGACTTTGAACCAAGAGATCTGCTTACCCAGCTCTCCTGGAATTA 303  
QY 240 AAGGCGTGTACTACTATTTGCTGGACCTAAGATTTTCATGATCAGTATGCTTCAAGATCT 299  
Db 304 AAGGCGTGTACTAC-CTTGCTGGACCTAAGATTTTCATGATCAGTATGCTTCAAGATCT 362  
QY 300 CCATGTCACCAAGATCTCCATGTCAGATCCAAAGTCCAAAGTCCAAAGTCCATCCTCAAG 359  
Db 363 CCATGTCACCAAGATCTCCATGTCAGATCCAAAGTCCAAAGTCCAAAGTCCATCCTCAAG 422  
QY 360 ATCTGATCAG 419  
Db 423 ATCTGATCAG 482  
QY 420 TGTCTCAACCGTGTGCTGCTGTTTGGGAATATGTCAACAG----- 461  
Db 483 TGTCTCAACCGTGTGCTGCTGTTTGGGAATATGTCAACAGCGGCTCTTCTTT 542  
QY 462 -----AATCCAGAGTGTGTGAGAACAGATGGCAGAGGAGTGGTG 503  
Db 543 GTGGATATATCACACAAAAATCCAGAGTGTGTGAGAACAGATGGCAGAGGAGTGGTG 602  
QY 504 GTTCCCAAGCTGTTTAAAAATGGGACCCACAGATTATCAGAGAGACACAGTGAAGGAGC 563  
Db 603 GTTCCCAAGCTGTTTAAAAATGGGACCCACAGTTATCAAGAGAGACACAGTGAAGGAGC 662  
QY 564 GAGAGAAAAAGGTTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAA 623  
Db 663 GAGAGAAAAAGGTTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAA 722  
QY 624 TCCAAAGAACCCCGGATGTTTGTGACAGTGAACCCCGTGGAGGCGCCGATGTGTGGGA 683  
Db 723 TCCAAAGAACCCCGGATGTTTGTGACAGTGAACCCCGTGGAGGCGCCGATGTGTGGGA 782  
QY 684 AGGACTATGACACAGCTGCTGAGAAAGTACTAGCCACACAGAACTCAGTGTGGG 743  
Db 783 AGGACTATGACACAGCTGCTGAGAAAGTACTAGCCACACAGAACTCAGTGTGGG 842  
QY 744 GCTGACAGTGTGCTGTGGAAAGTACATTGAGCATTTACTTGAAGACTTTCTGGAGTC 803  
Db 843 GCTGACAGTGTGCTGTGGAAAGTACATTGAGCATTTACTTGAAGACTTTCTGGAGTC 902  
QY 804 TGTGATGATGATCTGATGTTGGCCATCGGGTCATATTTTACGTGATGATGACGACAC 863  
Db 903 TGTGATGATGATCTGATGTTGGCCATCGGGTCATATTTTACGTGATGATGACGACAC 962  
QY 864 CTCCCGGATGCTGCTGACACCTGACCTCTACATTCCTTACAGTCTTTGAGATCAG 923  
Db 963 CTCCCGGATGCTGCTGACACCTGACCTCTACATTCCTTACAGTCTTTGAGATCAG 1022  
QY 924 GTCTGAGAGAGGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGGACAT 983  
Db 1023 GTCTGAGAGAGGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGGACAT 1082  
QY 984 CTGGGCCCCATCCAGACAGAGTGCAGTTCCTCTTCTGATGAGAGTGAAGTCAAGTCTT 1043  
Db 1083 CTGGGCCCCATCCAGACAGGAGTGCAGTTCCTCTTCTGATGAGAGTGAAGTCAAGTCTT 1142  
QY 1044 TCAAGACAACTTCGGGGTGAACCTCTGGCCAGCTGGTAGCACAGCTCCAGGCGTGGTG 1103  
Db 1143 TCAAGACAACTTCGGGGTGAACCTCTGGCCAGCTGGTAGCACAGCTCCAGGCGTGGTG 1202  
QY 1104 GTACAAAGGCGAGTCCGAGAGTTCACCTATGAGAGCGGGAACTGTGCGCGCGGTACAT 1163  
Db 1203 GTACAAAGGCGAGTCCGAGAGTTCACCTATGAGAGCGGGAACTGTGCGCGCGGTACAT 1262  
QY 1164 TCCATTTGGAGAGGGGATTTTACTACACAGCGGGGATTTTGGAGGAAAGCGCTACTCA 1223  
Db 1263 TCCATTTGGAGAGGGGATTTTACTACACAGCGGGGATTTTGGAGGAAAGCGCTACTCA 1322  
QY 1224 CATTCTCAACCTCAGGAGGAGTCTTAAAGGGGATCCTCCAGGACAGAAAGATGACAT 1283

Db 1323 CATTCTCAACCTCAGGAGGAGTCTTTAAGGGGATCCTCCAGGACAGAAACATGACAT 1382  
QY 1284 AGAAGCCAGTGGGATGATGAGAGCCACCTCAACAAATACTTCTCTTCAACAAACCCAC 1343  
Db 1383 AGAAGCCAGTGGGATGATGAGAGCCACCTCAACAAATACTTCTCTTCAACAAACCCAC 1442  
QY 1344 TAAATCCTATCTCCAGAGTATTTGCTGGACTATCAGATAGGCTGCTTCCAGATATTTAA 1403  
Db 1443 TAAATCCTATCTCCAGAGTATTTGCTGGACTATCAGATAGGCTGCTTCCAGATATTTAA 1502  
QY 1404 AAGTGTCAAGGTAGCTTGGCAGACAGAAAAGATATAATTTGGTTAGAAATATGCTGACT 1463  
Db 1503 AAGTGTCAAGGTAGCTTGGCAGACAGAAAAGATATAATTTGGTTAGAAATATGCTGACT 1562  
QY 1464 TCAATTTGTGATGAAACTTGCACACTATTTCT 1495  
Db 1563 TCAATTTGTGATGAAACTTGCACACTATTTACT 1594

RESULT 11

US-08-213-200A-3  
: Sequence 3, Application US/08213200A  
: GENERAL INFORMATION:  
: APPLICANT: Galili, Dr. Uri  
: TITLE OF INVENTION: REPLIC, DR. PATRICIA M.  
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR VACCINES  
: TITLE OF INVENTION: COMPRISING ALPHA-GALACTOSYL EPITOPES  
: NUMBER OF SEQUENCES: 3  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
: STREET: Suite 1800, Two Penn Center Plaza  
: CITY: Philadelphia  
: STATE: PA  
: COUNTRY: USA  
: ZIP: 19102  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE:  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/213, 200A  
: FILING DATE: 15-MAR-1994  
: CLASSIFICATION: 424  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Daniel A. Monaco  
: REGISTRATION NUMBER: 30,480  
: REFERENCE/DOCKET NUMBER: 8760-2 US  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (215) 568-8383  
: TELEFAX: (215) 568-5549  
: INFORMATION FOR SEQ ID NO: 3:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1131 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: cdna  
US-08-213-200A-3

Query Match 67.1%; Score 1005.8; DB 6; Length 1131;  
Best Local Similarity 95.3%; Pred. No. 9e-284;  
Matches 1078; Conservative 0; Mismatches 2; Indels 51; Gaps 2;

QY 382 ATGAATGTCAGGAGAAAGTAATCTGTTGATGCTGATTTGCTCAACCGTGGTTGCTG 441  
Db 1 ATGAATGTCAGGAGAAAGTATCTGTTGATGCTGATTTGCTCAACCGTGGTTGCTG 60  
QY 442 TTTTGGGAATATGTCACAG-----AATT 465  
Db 61 TTTTGGGAATATGTCACAGCCCCAGAGCGCTCTTCTTGTGGATATATCACACAAATTT 120

```
QY 466 CCAGAGGTTGGTGAGAACAGATGGCAGAGAGGACTGGTGGTTCCCAAGCTGGTTTAAAAAT 525
Db 121 CCAGAGGTTGGTGAGAACAGATGGCAGAGAGGACTGGTGGTTCCCAAGCTGGTTTAAAAAT 180
QY 526 GGGACCCACAGTTATCAAGAACACACCTAGAGAGCGGAGAGAGAAAGGTTAGAAATGGA 585
Db 181 GGGACCCACAGTTATCAAGAACACACCTAGAGAGCGGAGAGAGAAAGGTTAGAAATGGA 240
QY 586 GATCGCATTTGA-----AGAGCCTCAGCTATGGGACTATGGGACTGGTTCAATCCAAAG 630
Db 241 GATCGCATTTGAGGAAGATGACACACAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAG 300
QY 631 AACCGCCCGGATGTTTGGACAGTACCCCGTGGAGAGCGCCGATTTGTGTGGGAAGCACT 690
Db 301 AACCGCCCGGATGTTTGGACAGTACCCCGTGGAGAGCGCCGATTTGTGTGGGAAGCACT 360
QY 691 TATGACACAGCTCTGCTGGAAGTACTACGCCACACAGAACTCACTCTGGGGCTGACA 750
Db 361 TATGACACAGCTCTGCTGGAAGTACTACGCCACACAGAACTCACTCTGGGGCTGACA 420
QY 751 GTGTTTGTCTGTGGGAAAGTACATTGAGCATTTACTTGAAGACTTTCTGGAGTCTGCTGAC 810
Db 421 GTGTTTGTCTGTGGGAAAGTACATTGAGCATTTACTTGAAGACTTTCTGGAGTCTGCTGAC 480
QY 811 ATGTACTTTCATGGTTGGCCATCGGCTCATATTTTACGTGATGATAGACACACCTCCCGG 870
Db 481 ATGTACTTTCATGGTTGGCCATCGGCTCATATTTTACGTGATGATAGACACACCTCCCGG 540
QY 871 ATGCTGTGCTGACCTGAACCTCTACATTTCTTACAAGTCTTTGAGATCAGGTCGTGAG 930
Db 541 ATGCTGTGCTGACCTGAACCTCTACATTTCTTACAAGTCTTTGAGATCAGGTCGTGAG 600
QY 931 AAGAGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGCACATCTCTGGCC 990
Db 601 AAGAGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGCACATCTCTGGCC 560
QY 991 CACATCCAGCAGAGTGCAGCTTCCTCTCTCTGATGAGAGTGCATCAAGTCTTTCAAGAC 1050
Db 661 CACATCCAGCAGAGTGCAGCTTCCTCTCTCTGATGAGAGTGCATCAAGTCTTTCAAGAC 720
QY 1051 AACTTCGGGGTGGAACTCTGGCCAGCTGGTAGCACAGCTCCAGCCTGGTGGTACAAG 1110
Db 721 AACTTCGGGGTGGAACTCTGGCCAGCTGGTAGCACAGCTCCAGCCTGGTGGTACAAG 780
QY 1111 GCCAGTCCCGAGAAGTTACCTATGAGAGCGGGAACCTGTCGCGCGCGTACATTCATTTC 1170
Db 781 GCCAGTCCCGAGAAGTTACCTATGAGAGCGGGAACCTGTCGCGCGCGTACATTCATTTC 840
QY 1171 GGAGAGGGGATTTTACTACACCGGGCCATTTTGGAGGAACGCTACTCACATTCCTC 1230
Db 841 GGAGAGGGGATTTTACTACACCGGGCCATTTTGGAGGAACGCTACTCACATTCCTC 900
QY 1231 AACCTCACCAGGAGTGCTTTAAGGGATCTCCAGGACAAGAACATCAGATAGAGCC 1290
Db 901 AACCTCACCAGGAGTGCTTTAAGGGATCTCCAGGACAAGAACATCAGATAGAGCC 960
QY 1291 CAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTCTTTTCAACAAACCCACTAAAAATC 1350
Db 961 CAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTCTTTTCAACAAACCCACTAAAAATC 1020
QY 1351 CTATCTCCAGATATTGCTGGGACTATCAGATAGGCGCTCCCTTCAGATATTTAAAGTGTCTC 1410
Db 1021 CTATCTCCAGATATTGCTGGGACTATCAGATAGGCGCTCCCTTCAGATATTTAAAGTGTCTC 1080
QY 1411 AAGGTAGCTTGGCAGACAAAGATATAATTTGGTTAGAAATAATGCTCTGA 1461
Db 1081 AAGGTAGCTTGGCAGACAAAGATATAATTTGGTTAGAAATAATGCTCTGA 1131
```

RESULT 12

US-09-173-270-3

; Sequence 3, Application US/09173270

; GENERAL INFORMATION:

```
;
; APPLICANT: GALILI, URI
; REPIK, PATRICIA M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR VACCINES
; COMPRISING ALPHA-GALACTOSYL EPITOPES
;
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/173,270
; FILING DATE: 23-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/704,548
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8760-2 C11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 3:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1131 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-173-270-3
```

```
Query Match 67.1%; Score 1005.8; DB 15; Length 1131;
Best Local Similarity 95.3%; Pred. No. 9e-284;
Matches 1078; Conservative 0; Mismatches 2; Indels 51; Gaps 2;

QY 382 ATCAATGTCAAGGAAAGTAATCCTGTTGATGCTGCTCAACCCGTGGTTCGCG 441
Db 1 ATCAATGTCAAGGAAAGTAATCCTGTTGATGCTGCTCAACCCGTGGTTCGCG 60
QY 442 TTTTGGGAATATGTCAACAG-----AATT 465
Db 61 TTTTGGGAATATGTCAACAGCCGCTCTTCTTGGATATATCACACAAATTT 120
QY 466 CCAGAGGTTGGTGAGAACAGATGGCAGAGGACTGGTGGTTCCCAAGCTGGTTTAAAAAT 525
Db 121 CCAGAGGTTGGTGAGAACAGATGGCAGAGGACTGGTGGTTCCCAAGCTGGTTTAAAAAT 180
QY 526 GGGACCCACAGTTATCAAGAACACACCTAGAGAGCGGAGAGAGAAAGGTTAGAAATGGA 585
Db 181 GGGACCCACAGTTATCAAGAACACACCTAGAGAGCGGAGAGAGAAAGGTTAGAAATGGA 240
QY 586 GATCGCATTTGA-----AGAGCCTCAGCTATGGGACTGGTTCAATCCAAAG 630
Db 241 GATCGCATTTGAGGAAGATGACACACAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAG 300
QY 631 AACCGCCCGGATGTTTGGACAGTACCCCGTGGAGAGCGCCGATTTGTGTGGGAAGCACT 690
Db 301 AACCGCCCGGATGTTTGGACAGTACCCCGTGGAGAGCGCCGATTTGTGTGGGAAGCACT 360
QY 691 TATGACACAGCTCTGCTGGAAGTACTACGCCACACAGAACTCACTCTGGGGCTGACA 750
Db 361 TATGACACAGCTCTGCTGGAAGTACTACGCCACACAGAACTCACTCTGGGGCTGACA 420
```

[illegible]

RESULT 13

US-08-214-580-2

; Sequence 2, A

: GENERAL INFORMATION:

APPLICANT: Sandrin, Mauro S

APPLICANT: MCKENZIE, IAN F. C.

**TITLE OF INVENTION: XENOTRANSPLANTATION OF HUMAN EMBRYONAL TISSUES**

TITLE OF INVENTION: TH

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Maurice M

STREET: 1951 Burr

; CITY: Fairfile

; STATE: Conn

; COUNTRY: USA

ZIP: 06430  
COMPUTED PRESSURE BODY:

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 5 inch

```

;; MEDIUM TYPE: 3.5 Inch, 750  
: COMPUTER: DELL 486/50

```

; ; COMPUTER: DELL 486/50
; ; OPERATING SYSTEM: DOS 6.3

```

```

; ; OPERATING SYSTEM: DOS 6.0
; ; SOFTWARE: Wordperfect 6.0

```

```

; SOFTWARE: wordperfect 6.0
; CURRENT APPLICATION DATA.

```

QY 951 CATGATCGCGCATGAAGACCATTTGGGAGACATCTCTGGGCCACATCCAGCACGAGTCTGA 1010  
Db 619 CATGATCGCGCATGAAGACCATTCGGGAGCAGATCTCTGGGCCACATCCAGCACGAGTCTGA 678  
QY 1011 CTTCCCTCTCTGATGAGCTGATCAAGTCTTTTCAAGACAACCTTGGGGTGGAACTCT 1070  
Db 679 CTTCCCTCTCTGATGAGCTGATCAAGTCTTTTCAAGACAACCTTGGGGTGGAGACCTT 738  
QY 1071 GGCCAGCTGTTAGCACAGCTCCAGGCGTGGTGTACAAGGCCAGTCCCGAGAGTTTCA 1130  
Db 739 GGCCAGCTGTTAGCACAGCTCCAGGCGTGGTGTACAAGGCCAGTCCCGAGAGTTTCA 798  
QY 1131 CTATGAGAGCGGGAAGTGTGCGCGCGGTACATTCATTCGGAGAGGGGATTTTACTA 1190  
Db 799 CTACGAGAGCGGGAAGTGTGCGCGCGGTACATTCATTCGGAGAGGGGATTTTACTA 858  
QY 1191 CCAGCGGCGCATTTTGGAGGAAGCGCTACTACATCTCAACCTCACCAGGAGTCTT 1250  
Db 859 CCAGCGGCGCATTTTGGGGAAGACCCACTCAGGTTCTAAACATCACTCAGGAGTCTT 918  
QY 1251 TAAGGGGATCCTCCAGGACAAGAAACATGACATAGAACCCAGTGGCATGATGAGAGCCA 1310  
Db 919 CAAGGAATCCTCCAGGACAAGAAATGACATAGAACCCAGTGGCATGATGAGAGCCA 978  
QY 1311 CCTCAACAATACTCTCTTTTCAACAAACCCACTAAATCTATCTCCAGAGATTTGCTG 1370  
Db 979 TCTAAACAAGTATTTCTCTTCAACAAACCCACTAAATCTATCTCCAGAGATTTGCTG 1038  
QY 1371 GGACTATCAGATAGGCTGCTTCAGATATTAAAGTGTCAAGTAGCTTGGCAGACAAA 1430  
Db 1039 GGATATATCATATAGGCTGCTGAGATATTAGGATTTCAAGATAGCTTGGCAGAAAA 1098  
QY 1431 AGAGTAAATTTGGTTAGAAATTAATGCTGACTTCAAAATTTGTATGGAAACTT 1483  
Db 1099 AGAGTAAATTTGGTTAGAAATTAATGCTGACTTCAAAATTTGTATGGAAACTT 1151

RESULT 14

US-08-260-201-5  
; Sequence 5, Application US/08260201  
; GENERAL INFORMATION:  
; APPLICANT: Sandrin, Mauro S.  
; APPLICANT: Fodor, William L.  
; APPLICANT: Squinto, Stephen P.  
; APPLICANT: McKenzie, Ian F. C.  
; TITLE OF INVENTION: Methods for Reducing  
; TITLE OF INVENTION: Hyperacute Rejection of Xenografts  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maurice M. Klee  
; STREET: 1951 Burr Street  
; CITY: Fairfield  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 06430  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 Inch, 750 kb storage  
; COMPUTER: Dell 486/50  
; OPERATING SYSTEM: DOS 6.2  
; SOFTWARE: WordPerfect 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/260,201  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Klee, Maurice M.  
; REGISTRATION NUMBER: 30,399  
; REFERENCE/DOCKET NUMBER: ALX-144  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (203) 255-1400

; TELEFAX: (203) 254-1101  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1423 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Double  
; TOPOLOGY: Linear  
; MOLECULE TYPE: cDNA to mRNA  
; DESCRIPTION: galactosyl transferase,  
; DESCRIPTION: full coding sequence  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Sus scrofa  
; US-08-260-201-5

Query Match 48.4%; Score 726.6; DB 6; Length 1423;  
Best Local Similarity 78.8%; Pred. No. 1.le-201;  
Matches 893; Conservative 0; Mismatches 234; Indels 6; Gaps 2;  
QY 354 CTCAGATCTGGATCACAGGAGAAATAATGAATGTCAAGSGAAAGCTAATCTCTGTTGAT 413  
Db 22 CCAGCTTCTGCCGATCAGGAGAAATAATGAATGTCAAAGAGAGAGTGGTCTCTGTCAT 81  
QY 414 GCTGATTTGCTCAACCGTGTCTGCTGTTTGGGAATATCTCAACAGAAATTCACAGAGT 473  
Db 82 GCTGCTTGTCTCAACTGTAATGTTGTTTGGGAATACATCAACAGAAACCCAGAAGT 141  
QY 474 TGTGAGAACAGATGGGAGAGGACTGGTGGTCCCAAGCTGGTTTAAAATGGGACCCA 533  
Db 142 TGG---CAGCAGTGTCTCAGAGGGGCTGGTGGTTCGAGAGTGGTTTAAACAAATGGGACTCA 198  
QY 534 CAGTTATCAAGAAGACAACGCTAGAAGGACGAGAGAAAGS---GTAGAAATGGAGATCG 590  
Db 199 CAGTTACCAGAGAGAGAGAGAGCTATAGCAAGAGAAAGAGAGAGAGAGAGAGAGAG 258  
QY 591 CATTGAAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAGAACCGCCCGGATGTTTGGAC 650  
Db 259 CAGAGGAGAGCTTCCGCTAGTGGACTGGTTTAACTCTGAGAAACGCCAGAGTCTGAG 318  
QY 651 AGTACCCCGTGGAAAGCGCGGATTTGTCGGGAAGGCACTTATGACACAGCTCTGCTGGA 710  
Db 319 CATAAACAGATGGAAGGCTCCAGTGGTATGGGAAGGCACTTACAAAGAGAGCCGCTTAGA 378  
QY 711 AAAGTACTACGCACACAGAGAACTCAGTGTGGGCTCAGAGTGTCTGCTGGGAAAGTA 770  
Db 379 TAATTATTGCCAAACAGAGAAATACCGTGGGCTTACCGGTTTCTGCTGTCGGAAGATA 438  
QY 771 CATTGAGCATTTACTTGAAGACTTTCTGGAGTCTGTGACATGTACTTCTATGTTGGCCA 830  
Db 439 CATTGAGCATTTACTTGGAGGAGTCTTAAATATCTGCAAAATACATCTTATGTTGGCCA 498  
QY 831 TCGGGTCATATTTTACGTCATGATAGACAGACACCTCCCGGATGCCCTGCTGCACCTGAA 890  
Db 499 CAAAGTCATCTTTTACATCATGTTGGATGATATCTCCAGGATGCCCTTGAATAGAGTGG 558  
QY 891 CCCTCTACATTTCTTACAAGTCTTTGAGATCAGGCTCTGAGAAGAGTGGCAGGATATCAG 950  
Db 559 TCCTCTCGGTTCTTTTAAAGTGTGTTGAGATCAAGTCCGAGAAAGAGTGGCAAGACATCAG 618  
QY 951 CATGATCGCATGAAGACCATTTGGGAGCACATCTCTGGCCACATCCAGCACAGAGTCTGA 1010  
Db 619 CATGATCGCATGAAGACCATTCGGGAGCACATCTCTGGCCACATCCAGCACAGAGTCTGA 678  
QY 1011 CTTCTCTCTTCTGATGGAGCTGGATCAAGTCTTTTCAAGACAACCTTCGGGTGGAACTCT 1070  
Db 679 CTTCTCTCTTCTGATGGAGCTGGATCAAGTCTTTTCAAGACAACCTTCGGGTGGAGACCT 738  
QY 1071 GGCCAGCTGTAGCACAGCTCAGGCGTGGTGTGTAAGGCCAGTCCCGAGAGAGTTTCA 1130  
Db 739 GGCCAGCTGTAGCACAGCTCAGGCGTGGTGTGTAAGGCCAGTCCCGAGAGAGTTTCA 929

Qy 1131 CTATGAGAGCGGGAAGTCTCGCGCGGTACATATCCATTCGGAGAGGGGATTTTACTA 1190  
Db 799 CTACGAGAGCGGAAGAGTCCGACGCTACATTCCTGTTGGCCAGGGGATTTTATTA 858  
Qy 1191 CCAGCGGCCATTTTGGAGGAGCGCTACTACATCTCTCAACCTCACCAGGGAGTCTT 1250  
Db 859 CCAGCGAGCCATTTTGGGGAACACCCACTCAGGTCTTAAACATCACTCAGGAGTCTT 918  
Qy 1251 TAAGGGATCTCCAGGACAAGAAACATGACATAGAAGCCAGTGGCATGATGAGAGCA 1310  
Db 919 CAAGGGAATCTCCAGGACAAGAAATGACATAGAAGCCAGTGGCATGATGAGAGCA 978  
Qy 1311 CCTCAACAAATCTCTCTTTTCAACAAACCCACTAAATCTCTCCAGAGTATGCTG 1370  
Db 979 TCTAAACAAGTATTTCTCTTCAACAAACCCACTAAATCTCTATCCCGAGAAATCTGCTG 1038  
Qy 1371 GGACTATCAGATAGCGCTTCCCTTCCAGATATTAAGTCTCAAGGTAGCTGGCAGACAA 1430  
Db 1039 GGATTATCATATAGCATGTCTGGGATATTTAGGATTTCTCAAGATAGCTTGGCAGAAAA 1098  
Qy 1431 AGATATAAATTTGGTTAGAAATATGTCTGACTTCAAAATTTGTATGGAAACTT 1483  
Db 1099 AGATATAAATTTGGTTAGAAATATCATCTGACTTTAAATTTGTGCGCAGCTTT 1151

RESULT 15

US-08-278-282-5  
Sequence 5, Application US/08278282  
GENERAL INFORMATION:  
APPLICANT: Sandrin, Mauro S.  
APPLICANT: Fodor, William L.  
APPLICANT: Rother, Russell P.  
APPLICANT: Suinto, Stephen P.  
APPLICANT: McKenzie, Ian F. C.  
TITLE OF INVENTION: Methods for Reducing  
TITLE OF INVENTION: Hyperacute Rejection of Xenografts  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maurice M. Klee  
STREET: 1951 Burr Street  
CITY: Fairfield  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06430  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 750 Kb storage  
COMPUTER: Dell 486/50  
OPERATING SYSTEM: DOS 6.2  
SOFTWARE: WordPerfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/278,282  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/260,201  
FILING DATE: June 15, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: ALX-144.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255-1400  
FAX: (203) 254-1101  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1423 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
DESCRIPTION: galactosyl transferase,  
DESCRIPTION: full coding sequence  
HYPOTHETICAL: NO

; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Sus scrofa  
US-08-278-282-5  
Query Match 48.4%; Score 726.6; DB 6; Length 1423;  
Best Local Similarity 78.8%; Pred. No. 1.1e-201;  
Matches 893; Conservative 0; Mismatches 234; Indels 6; Gaps 2;  
Qy 354 CTCAAGATCTGGATCAGAGGAGAAATTAATGTCAAGGAGAAAGTAATCTCTGTTGAT 413  
Db 22 CCCAGCTTCTGCCCATCAGGAGAAATTAATGAATGTCAAGGAGAGTGGTCTGTCTCAAT 81  
Qy 414 GCTGATTTCTCAACCGTGGTGTGTTGGGAATATCTCAACAGAAATTCAGAGGT 473  
Db 82 GCTGCTTCTCACTGTATGGTGTGTTGGGAATATCATCAACAAACCAGAGT 141  
Qy 474 TGGTGAACAGATGGCAGAGAGTGGTGTCTCCAGAGTGGTGTAAATGGAGCCCA 533  
Db 142 TGG--CAGCAGTGTCTCAGAGGGGTGGTGTCTCCAGAGTGGTGTAAACATGGAGCTCA 198  
Qy 534 CAGTTATCAAGAACACACACCTAGAGGACGAGAGAGAGG--GTAGAAATGGAGATCG 590  
Db 199 CAGTTATCAAGAACAGAGAGAGCTATAGGCAAGAAAGAAAGAAAGAAAGAGCA 258  
Qy 591 CATTTGAAGAGCTCAGCTATGGGACTGTGTTCAATCCAAAGAACCGCCCGGATGTTTTCAC 650  
Db 259 CAGAGGAGAGCTTCCGCTAGTGGAGTGTGTTTAACTCTCAGAAACGCCAGAGGTCTGTGAC 318  
Qy 651 AGTGACCCCGTGGAGGCGCGATTTGTGGGAAGGCACTTATGACACAGCTCTGCTGGA 710  
Db 319 CATAAACAGATGGAAGGCTCCAGTGGTATGGGAAGGCACTTACAACAGAGCGCTTAGA 378  
Qy 711 AAAGTACTACCCACACAGAACTCACTGTGGGCTGACAGTGTGCTGTGGGAAGTA 770  
Db 379 TAATTTATGCCAAACAGAAAAATACCGTGGGCTTTCAGGCTTTTGGTCTGGGAAGATA 438  
Qy 771 CATTTGAGCATTTAGAAAGCTTTCTGGAGTCTGCTGACATGTACTCATCTGTTGGGCA 830  
Db 439 CATTTGAGCATTTAGAAAGCTTTCTGGAGGATTTCTTAATATCTGCAATATACATCTTTCAT 498  
Qy 831 TCGGGTCATATTTTACGTCATGATAGAGACACCTCCCGGATGGCTGTGCTGACCTGAA 890  
Db 499 CAAAGTCATCTTTTACATCATGCTGGTGTATCTCCAGGATGGCTTTGATAGAGCTGG 558  
Qy 891 CCCTTACATTTCTTACAAGCTTTTGAAGTCTGAGAGAGGCTGGCAGGATATCAG 950  
Db 559 TCCTCTCGTTCCTTTAAAGTGTGTTGAGATCAAGTCCGAGAGAGGTTGGCAAGACATCAG 618  
Qy 951 CATGATCGCATGAAGACCATTTGGGAGGACATCTTGGCCACATCCAGCAGAGGTGCA 1010  
Db 619 CATGATCGCATGAAGACCATTTGGGAGGACATCTTGGCCACATCCAGCAGAGGTGCA 678  
Qy 1011 CTTCCTCTTCTGATGGAGTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAACTCT 1070  
Db 679 CTTCCTCTTCTGATTTGACGTGGATCAGGTCTTCCAAAACAACCTTTGGGGTGGAGACCT 738  
Qy 1071 GGGCAGCTGGTACACAGCTCCAGGCTGGTGTACAAGGCCAGTCCCGAGAAAGTTTCA 1130  
Db 739 GGGCAGCTGGTGGCTCAGCTACAGGCTGGTGTACAGGACATCTCTGACAGAGTTCAC 798  
Qy 1131 CTATGAGAGCGGGAACCTGTCGGCGGCTACATTTCCATTCGAGAGGGGGATTTTTACTA 1190  
Db 799 CTACGAGAGCGGGAAGGAGTCCGACGCTACATTTCCGTTTGGCCAGGGGATTTTTATTA 858  
Qy 1191 CCAGCGGCCATTTTGGAGGAAGCCCTACTTCAATCTCAACCTCACCAGGGAGTCTT 1250  
Db 859 CCAGCGGACCATTTTGGGGAACACCCACTCAGGTCTTAAACATCACTCAGGAGTCTT 918  
Qy 1251 TAAGGGATCTCTCAGGACAAGAAACATGACATAGAAGCCCGGCTGGCATGATGAGAGCA 1310  
Db 919 CAAGGGAATCTCCAGGACAAGAAATGACATAGAAGCCCGGCTGGCATGATGAGAGCA 978



QY 1311 CCTCAACAATACTTCCTTTTCAACAAACCCACTAAAAATCCTATCTCCAGAGTATTGCTG 1370  
  || |||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
Db 979 TCTAAACAAGTATTTCCTTCTCAACAACCCACTAAAAATCTTATCCCCAGAACTACTGCTG 1038  
  || |||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
QY 1371 GGACTATCAGATAGGCTGCCCTTCAGATATTAAAAAGTGTCAAGGTAGCTTGGCAGACAAA 1430  
  || |||| || |||| || || |||| || |||| || |||| || |||| || |||| || ||||  
Db 1039 GGATTATCATATAGGCATGCTGTGGATATTAGGATTGTCAAGATAGCTTGGCAGAAAAA 1098  
  || |||| || |||| || || |||| || |||| || |||| || |||| || |||| || ||||  
QY 1431 AGAGTATAATTTGGTTAGAAAATAATGCTGACTTCAAAATTTGATGGAACCTT 1483  
  || |||| || |||| || || |||| || |||| || |||| || |||| || |||| || ||||  
Db 1099 AGAGTATAATTTGGTTAGAAAATAACATCTGACTTTAAATTTGCCAGCAGTTT 1151  
  || |||| || |||| || || |||| || |||| || |||| || |||| || |||| || ||||

Search completed: May 17, 2002, 20:30:28  
Job time: 9097 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:05:57 ; Search time 109.08 Seconds  
(without alignments)  
1268.739 Million cell updates/sec

Title: US-09-863-475A-4  
Perfect score: 2136  
Sequence: 1 MITMLQDLHVKNKISMSRSKS.....IKSVKVAQOTKEYNLVRNVV 394

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3516493 seqs, 351254056 residues  
Total number of hits satisfying chosen parameters: 3516493

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep:\*  
25: /cgn2\_6/ptodata/2/paa/US60\_MERGED\_COMB.pep1:\*  
26: /cgn2\_6/ptodata/2/paa/US10\_MERGED\_COMB.pep1:\*  
27: /cgn2\_6/ptodata/2/paa/US09\_MERGED\_COMB.pep2:\*  
28: /cgn2\_6/ptodata/2/paa/US09\_MERGED\_COMB.pep2:\*  
29: /cgn2\_6/ptodata/2/paa/US08\_MERGED\_COMB.pep1:\*  
30: /cgn2\_6/ptodata/2/paa/US07\_MERGED\_COMB.pep1:\*  
31: /cgn2\_6/ptodata/2/paa/US06\_MERGED\_COMB.pep1:\*  
32: /cgn2\_6/ptodata/2/paa/PCT\_MERGED\_COMB.pep1:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	2136	100.0	394	7 US-08-379-040-2 Sequence 2, Appl1

2	2136	100.0	394	12 US-08-823-489-4 Sequence 4, Appl1
3	2136	100.0	394	19 US-09-593-316-10 Sequence 10, Appl1
4	2136	100.0	394	22 US-09-863-475-4 Sequence 4, Appl1
5	2136	100.0	394	27 US-09-863-475A-4 Sequence 4, Appl1
6	2136	100.0	394	27 US-09-946-034-2 Sequence 2, Appl1
7	1969	92.2	359	28 US-09-994-427A-11 Sequence 11, Appl1
8	1950	91.3	371	5 US-08-188-607B-12 Sequence 12, Appl1
9	1950	91.3	371	7 US-08-378-617-12 Sequence 12, Appl1
10	1950	91.3	371	13 US-08-984-900-12 Sequence 14, Appl1
11	1950	91.3	371	28 US-09-870-759-74 Sequence 72, Appl1
12	1542.5	72.2	376	28 US-09-994-427A-12 Sequence 12, Appl1
13	1539.5	72.1	376	6 US-08-213-200A-2 Sequence 2, Appl1
14	1539.5	72.1	376	19 US-09-593-316-6 Sequence 6, Appl1
15	1539.5	72.1	376	27 US-09-173-270-2 Sequence 2, Appl1
16	1539.5	72.1	376	28 US-09-994-427A-6 Sequence 6, Appl1
17	1539.5	72.1	376	28 US-09-995-419A-9 Sequence 9, Appl1
18	1532	71.7	363	16 US-09-230-091-4 Sequence 4, Appl1
19	1523.5	71.3	376	28 US-09-994-427A-13 Sequence 13, Appl1
20	1513	70.8	375	16 US-09-230-091-2 Sequence 2, Appl1
21	1510	70.7	371	13 US-08-984-900-10 Sequence 10, Appl1
22	1510	70.7	371	19 US-09-593-316-8 Sequence 8, Appl1
23	1510	70.7	371	28 US-09-994-427A-10 Sequence 10, Appl1
24	1493	69.9	371	5 US-08-188-607B-10 Sequence 10, Appl1
25	1476.5	69.1	368	5 US-08-188-607B-11 Sequence 10, Appl1
26	1476.5	69.1	368	7 US-08-378-617-11 Sequence 11, Appl1
27	1476.5	69.1	368	13 US-08-984-900-11 Sequence 11, Appl1
28	1476.5	69.1	368	19 US-09-593-316-4 Sequence 4, Appl1
29	1476.5	69.1	368	28 US-09-994-427A-9 Sequence 9, Appl1
30	1476.5	69.1	369	19 US-09-593-316-2 Sequence 2, Appl1
31	1471	68.9	369	28 US-09-994-427A-8 Sequence 8, Appl1
32	1471	68.9	369	28 US-09-995-419A-7 Sequence 7, Appl1
33	1471	68.9	354	16 US-09-230-091-6 Sequence 6, Appl1
34	1431.5	67.0	342	16 US-09-230-091-8 Sequence 8, Appl1
35	1428.5	66.9	342	27 US-09-994-427A-7 Sequence 7, Appl1
36	945	44.2	227	28 US-09-994-427A-3 Sequence 3, Appl1
37	690.5	32.3	335	24 US-60-178-574-9 Sequence 9, Appl1
38	688	32.2	353	28 US-09-994-427A-3 Sequence 3, Appl1
39	688	32.2	353	28 US-09-995-419A-11 Sequence 11, Appl1
40	687.5	32.2	354	28 US-09-994-427A-5 Sequence 5, Appl1
41	687.5	32.2	354	28 US-09-995-419A-13 Sequence 13, Appl1
42	673	31.5	347	18 US-09-488-725A-2396 Sequence 2396, Ap
43	646	30.2	402	18 US-09-488-725A-5968 Sequence 5968, Ap
44	632	29.6	197	19 US-09-593-316-12 Sequence 12, Appl1
45	482	22.6	211	27 US-09-611-526-3056 Sequence 3056, Ap

ALIGNMENTS

RESULT 1  
US-08-379-040-2  
; Sequence 2, Application US/08379040  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, David K.C.  
; APPLICANT: Koren, Eugen  
; TITLE OF INVENTION: GENETICALLY ENGINEERED ANIMALS FOR USE  
; NUMBER OF INVENTION: AS ORGAN DONORS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 1100 Peachtree Street, Suite 2800  
; CITY: Atlanta  
; STATE: GA  
; COUNTRY: USA  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/379,040  
; FILING DATE:

CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,817  
FILING DATE: 20-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: BMC100  
TELEPHONE: (404)-815-6508  
TELEFAX: (404)-815-6555  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORGANISM: Mus musculus  
US-08-379-040-2

Query Match 100.0%; Score 2136; DB 7; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.7e-212;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSTSLPSSRSQSEKIMNVKGVILLMLLIVSTVVVWFVEYV 60  
DB 1 MITMLQDLHVNKISMSRSKSTSLPSSRSQSEKIMNVKGVILLMLLIVSTVVVWFVEYV 60  
QY 61 NRIPEVGENRWQKDWPFSPFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLWDWPNKRN 120  
DB 61 NRIPEVGENRWQKDWPFSPFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLWDWPNKRN 120  
QY 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTGVTFAVGKYEIHYLEDLFLESADMY 180  
DB 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTGVTFAVGKYEIHYLEDLFLESADMY 180  
QY 181 FMVGHVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
DB 181 FMVGHVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
QY 241 QHEVDLFCHMDVQVDFQDNFVETLGOLVAQLQAWMYKASPEKTYERRELSAAYIPFGE 300  
DB 241 QHEVDLFCHMDVQVDFQDNFVETLGOLVAQLQAWMYKASPEKTYERRELSAAYIPFGE 300  
QY 301 GDFYHAAIFGGTPTHILNLTRECFKGLQDKKHDI EAQWHDHSHLNKYFLFNKPTKILS 360  
DB 301 GDFYHAAIFGGTPTHILNLTRECFKGLQDKKHDI EAQWHDHSHLNKYFLFNKPTKILS 360  
QY 361 PEYCDYQIGLPSDIKSVKVAQTKYENLVNRYV 394  
DB 361 PEYCDYQIGLPSDIKSVKVAQTKYENLVNRYV 394

## RESULT 2

US-08-823-489-4  
Sequence 4. Application US/08823489  
GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia

COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/823,489  
FILING DATE: 25-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-823-489-4

Query Match 100.0%; Score 2136; DB 12; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.7e-212;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSTSLPSSRSQSEKIMNVKGVILLMLLIVSTVVVWFVEYV 60  
DB 1 MITMLQDLHVNKISMSRSKSTSLPSSRSQSEKIMNVKGVILLMLLIVSTVVVWFVEYV 60  
QY 61 NRIPEVGENRWQKDWPFSPFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLWDWPNKRN 120  
DB 61 NRIPEVGENRWQKDWPFSPFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLWDWPNKRN 120  
QY 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTGVTFAVGKYEIHYLEDLFLESADMY 180  
DB 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTGVTFAVGKYEIHYLEDLFLESADMY 180  
QY 181 FMVGHVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
DB 181 FMVGHVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
QY 241 QHEVDLFCHMDVQVDFQDNFVETLGOLVAQLQAWMYKASPEKTYERRELSAAYIPFGE 300  
DB 241 QHEVDLFCHMDVQVDFQDNFVETLGOLVAQLQAWMYKASPEKTYERRELSAAYIPFGE 300  
QY 301 GDFYHAAIFGGTPTHILNLTRECFKGLQDKKHDI EAQWHDHSHLNKYFLFNKPTKILS 360  
DB 301 GDFYHAAIFGGTPTHILNLTRECFKGLQDKKHDI EAQWHDHSHLNKYFLFNKPTKILS 360  
QY 361 PEYCDYQIGLPSDIKSVKVAQTKYENLVNRYV 394  
DB 361 PEYCDYQIGLPSDIKSVKVAQTKYENLVNRYV 394

## RESULT 3

US-09-593-316-10  
Sequence 10. Application US/09593316  
GENERAL INFORMATION:  
APPLICANT: Clark Mr., John  
APPLICANT: Denning, Chris  
TITLE OF INVENTION: Animal Tissue For Xenotransplantation  
FILE REFERENCE: 730/002  
CURRENT APPLICATION NUMBER: US/09/593,316

; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: 60/204,148
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Musca sp.
US-09-593-316-10

Query Match 100.0%; Score 2136; DB 19; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.7e-212;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKINNVKGVILLMLLIVSTVVVWFVEYV 60
DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKINNVKGVILLMLLIVSTVVVWFVEYV 60

QY 61 NRIPVEGENRWQKDMWFPSPFKNGTHSYQEDNVGRRKGRNGDRIEEPQLWDFNPKNR 120
DB 61 NRIPVEGENRWQKDMWFPSPFKNGTHSYQEDNVGRRKGRNGDRIEEPQLWDFNPKNR 120

QY 121 PDVLTVPKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKYIEHYLEDPLESADMY 180
DB 121 PDVLTVPKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKYIEHYLEDPLESADMY 180

QY 181 FMVGHVIFVYVIMDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMRRKKTIGEHILAH 240
DB 181 FMVGHVIFVYVIMDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMRRKKTIGEHILAH 240

QY 241 QHEVDFLCMDVDQVQDNFGVETLGLVAQLQAWYKASPEKTYERRELSAAYIPFGE 300
DB 241 QHEVDFLCMDVDQVQDNFGVETLGLVAQLQAWYKASPEKTYERRELSAAYIPFGE 300

QY 301 GDFYHAAIFGGTPTTHLNLTRCFKGILODKKHIDEAOWHDESHLNKYFLFNKPTKILS 360
DB 301 GDFYHAAIFGGTPTTHLNLTRCFKGILODKKHIDEAOWHDESHLNKYFLFNKPTKILS 360

QY 361 PEYCWDYQIGLPSDIKSVKAWOTKEYNLVRNV 394
DB 361 PEYCWDYQIGLPSDIKSVKAWOTKEYNLVRNV 394

RESULT 4
US-09-863-475-4
; Sequence 4, Application US/09863475
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
;
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/863,475
; FILING DATE: 24-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/914,281
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-863-475-4

Query Match 100.0%; Score 2136; DB 22; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.7e-212;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKINNVKGVILLMLLIVSTVVVWFVEYV 60
DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKINNVKGVILLMLLIVSTVVVWFVEYV 60

QY 61 NRIPVEGENRWQKDMWFPSPFKNGTHSYQEDNVGRRKGRNGDRIEEPQLWDFNPKNR 120
DB 61 NRIPVEGENRWQKDMWFPSPFKNGTHSYQEDNVGRRKGRNGDRIEEPQLWDFNPKNR 120

QY 121 PDVLTVPKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKYIEHYLEDPLESADMY 180
DB 121 PDVLTVPKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKYIEHYLEDPLESADMY 180

QY 181 FMVGHVIFVYVIMDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMRRKKTIGEHILAH 240
DB 181 FMVGHVIFVYVIMDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMRRKKTIGEHILAH 240

QY 241 QHEVDFLCMDVDQVQDNFGVETLGLVAQLQAWYKASPEKTYERRELSAAYIPFGE 300
DB 241 QHEVDFLCMDVDQVQDNFGVETLGLVAQLQAWYKASPEKTYERRELSAAYIPFGE 300

QY 301 GDFYHAAIFGGTPTTHLNLTRCFKGILODKKHIDEAOWHDESHLNKYFLFNKPTKILS 360
DB 301 GDFYHAAIFGGTPTTHLNLTRCFKGILODKKHIDEAOWHDESHLNKYFLFNKPTKILS 360

QY 361 PEYCWDYQIGLPSDIKSVKAWOTKEYNLVRNV 394
DB 361 PEYCWDYQIGLPSDIKSVKAWOTKEYNLVRNV 394

RESULT 5
US-09-863-475A-4
; Sequence 4, Application US/09863475A
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
;
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/863,475
; FILING DATE: 24-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: 07/914,281  
FILING DATE: 20-JUL-1992  
CLASSIFICATION: <Unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-863-475A-4

Query Match 100.0%; Score 2136; DB 27; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.7e-212; Indels 0; Gaps 0;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MITLQDLHVNKISMSRSKSETSLPSSRSQSEKIMNVKGVILLMLLIVSTVVVWFVEYV 60  
Db 1 MITLQDLHVNKISMSRSKSETSLPSSRSQSEKIMNVKGVILLMLLIVSTVVVWFVEYV 60  
Qy 61 NRIPEVGENRQKDWPFQFNGVETLQGLVAQLQAWYKASPEKTYERRELSAAYIPGE 120  
Db 61 NRIPEVGENRQKDWPFQFNGVETLQGLVAQLQAWYKASPEKTYERRELSAAYIPGE 120  
Qy 121 PDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGTLVFAVGKVIHYLEDLFLESADMY 180  
Db 121 PDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGTLVFAVGKVIHYLEDLFLESADMY 180  
Qy 181 FMVGHRIYFYVMIDTTSRMPVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
Db 181 FMVGHRIYFYVMIDTTSRMPVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
Qy 241 QHEVDLFCEMDVDQVDFQDFGVTGLQGLVAQLQAWYKASPEKTYERRELSAAYIPGE 300  
Db 241 QHEVDLFCEMDVDQVDFQDFGVTGLQGLVAQLQAWYKASPEKTYERRELSAAYIPGE 300  
Qy 301 GDFYHAAIFGCTPTTHILNLTRECFKGLQDKKHIDIEAQWHDHSLNKYFLFNKPTKILS 360  
Db 301 GDFYHAAIFGCTPTTHILNLTRECFKGLQDKKHIDIEAQWHDHSLNKYFLFNKPTKILS 360  
Qy 361 PEYCDWQIGLPSDIKSVKVAWQTKYENLVRNV 394  
Db 361 PEYCDWQIGLPSDIKSVKVAWQTKYENLVRNV 394

RESULT 6  
US-09-946-034-2  
Sequence 2, Application US/09946034  
GENERAL INFORMATION:  
APPLICANT: Cooper, David K.C.  
TITLE OF INVENTION: GENETICALLY ENGINEERED ANIMALS FOR USE  
AS ORGAN DONORS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 1100 Peachtree Street, Suite 2800

CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-4530  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/946,034  
FILING DATE: 04-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,817  
FILING DATE: 20-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: BMC100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)-815-6508  
TELEFAX: (404)-815-6555  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-946-034-2

Query Match 100.0%; Score 2136; DB 27; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.7e-212; Indels 0; Gaps 0;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MITLQDLHVNKISMSRSKSETSLPSSRSQSEKIMNVKGVILLMLLIVSTVVVWFVEYV 60  
Db 1 MITLQDLHVNKISMSRSKSETSLPSSRSQSEKIMNVKGVILLMLLIVSTVVVWFVEYV 60  
Qy 61 NRIPEVGENRQKDWPFQFNGVETLQGLVAQLQAWYKASPEKTYERRELSAAYIPGE 120  
Db 61 NRIPEVGENRQKDWPFQFNGVETLQGLVAQLQAWYKASPEKTYERRELSAAYIPGE 120  
Qy 121 PDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGTLVFAVGKVIHYLEDLFLESADMY 180  
Db 121 PDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGTLVFAVGKVIHYLEDLFLESADMY 180  
Qy 181 FMVGHRIYFYVMIDTTSRMPVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
Db 181 FMVGHRIYFYVMIDTTSRMPVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
Qy 241 QHEVDLFCEMDVDQVDFQDFGVTGLQGLVAQLQAWYKASPEKTYERRELSAAYIPGE 300  
Db 241 QHEVDLFCEMDVDQVDFQDFGVTGLQGLVAQLQAWYKASPEKTYERRELSAAYIPGE 300  
Qy 301 GDFYHAAIFGCTPTTHILNLTRECFKGLQDKKHIDIEAQWHDHSLNKYFLFNKPTKILS 360  
Db 301 GDFYHAAIFGCTPTTHILNLTRECFKGLQDKKHIDIEAQWHDHSLNKYFLFNKPTKILS 360  
Qy 361 PEYCDWQIGLPSDIKSVKVAWQTKYENLVRNV 394  
Db 361 PEYCDWQIGLPSDIKSVKVAWQTKYENLVRNV 394

RESULT 7  
US-09-994-427A-11  
Sequence 11, Application US/09994427A

```

; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Schiffr, J. Michael
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE VECTORS FOR TREATING CANCER
; FILE REFERENCE: 083.002
; CURRENT APPLICATION NUMBER: US/09/994,427A
; CURRENT FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: 60/253,395
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-994-427A-11

Query Match      92.2%; Score 1969; DB 28; Length 359;
Best Local Similarity 100.0%; Pred. No. 3.5e-195;
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 MNVKGKILLMLIVSTVVVFWFVWVRIPEVGENRWQKDWFPKNGTHSYQEDNVEG 95
Db 1 MNVKGKILLMLIVSTVVVFWFVWVRIPEVGENRWQKDWFPKNGTHSYQEDNVEG 60
Qy 96 RREKGRNGRIEPPQLMDWPNKPNRPDLVTPWKAPVWEGTYDTALLEKYATOKLTV 155
Db 61 RREKGRNGRIEPPQLMDWPNKPNRPDLVTPWKAPVWEGTYDTALLEKYATOKLTV 120
Qy 156 GLTVFAVGKYIEHYLEDFLESADMYFMVGHVIFVYVIMIDTSTRMPVHLNPLHSLQVFEI 215
Db 121 GLTVFAVGKYIEHYLEDFLESADMYFMVGHVIFVYVIMIDTSTRMPVHLNPLHSLQVFEI 180
Qy 216 RSEKRWODISMWRKKTIGEHLAHIQHEVDLFECMDVDQVFOQNFVETLGLVLAQLOAW 275
Db 181 RSEKRWODISMWRKKTIGEHLAHIQHEVDLFECMDVDQVFOQNFVETLGLVLAQLOAW 240
Qy 276 WYKASPEKFTYERRELSAAYIPFEGDFYVYHAAIFGGTPTTHILNLTRECFKGLQKKHD 335
Db 241 WYKASPEKFTYERRELSAAYIPFEGDFYVYHAAIFGGTPTTHILNLTRECFKGLQKKHD 300
Qy 336 IEAQWHDHSHLNKYFLFNKPTKILSPYCWYDQIGLPSDIKSVKVAWQTKYENLVRNV 394
Db 301 IEAQWHDHSHLNKYFLFNKPTKILSPYCWYDQIGLPSDIKSVKVAWQTKYENLVRNV 359

RESULT      8
US-08-188-607B-12
; Sequence 12, Application US/08188607B
; GENERAL INFORMATION:
; APPLICANT: d'Apice, Anthony J.F.
; APPLICANT: Pearce, Martin J.
; APPLICANT: Robins, Allan J.
; APPLICANT: Crawford, Robert J.
; APPLICANT: Rathjen, Peter D.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF HYPERACUTE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patterson & Keough, P.A.
; STREET: 527 Marquette Avenue South
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh System 7.0
; SOFTWARE: Macintosh Text File
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,607B
; FILING DATE: 27-01-1994
```

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark S. Ellinger
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 1175.04-US-01
; TELEPHONE: 612/349-5740
; TELEFAX: 612/349-9266
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 371
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; HYPOTHEICAL: no
; ANTI-SENSE:
; FRAGMENT TYPE: Entire protein
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; STRAIN:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; FEATURE:
; NAME/KEY: Murine alpha-1,3-Galactosyltransferase amino acid sequence dedu
; PUBLICATION INFORMATION:
; AUTHORS: Joiasse, D. H.
; AUTHORS: Shaper, N.L.
; AUTHORS: Kim, D.
; AUTHORS: van den Eijnden, D.H.
; AUTHORS: Shaper, J. H.
; TITLE: Murine alpha-1,3-galactosyltransferase: A single gene locus specifi
; JOURNAL: The Journal of Biological Chemistry
; VOLUME: 267
; ISSUE: 8
; PAGES: 5534-5541
; DATE: 1992
US-08-188-607B-12

Query Match      91.3%; Score 1950; DB 5; Length 371;
Best Local Similarity 96.5%; Pred. No. 3.4e-193;
Matches 358; Conservative 1; Mismatches 0; Indels 12; Gaps 1;

Qy 36 MNVKGKILLMLIVSTVVVFWFVWVRIPEVGENRWQKDWFPKNGTHSYQEDNVEG 83
Db 1 MNVKGKILLMLIVSTVVVFWFVWVRIPEVGENRWQKDWFPKNGTHSYQEDNVEG 60
Qy 84 GTHSYQEDNVEGRREKGRNGRIEPPQLMDWPNKPNRPDLVTPWKAPVWEGTYDTAL 143
Db 61 GTHSYQEDNVEGRREKGRNGRIEPPQLMDWPNKPNRPDLVTPWKAPVWEGTYDTAL 120
Qy 144 LEKYATQKLTVGLTVFAVGKYIEHYLEDFLESADMYFMVGHVIFVYVIMIDTSTRMPV 203
Db 121 LEKYATQKLTVGLTVFAVGKYIEHYLEDFLESADMYFMVGHVIFVYVIMIDTSTRMPV 180
Qy 204 LNPVLSQVFEIRSEKRWQDISMRKKTIGEHLAHIQHEVDLFECMDVDQVFOQNFV 263
Db 181 LNPVLSQVFEIRSEKRWQDISMRKKTIGEHLAHIQHEVDLFECMDVDQVFOQNFV 240
Qy 264 TLGOLVAQLOAWYKASPEKFTYERRELSAAYIPFEGDFYVYHAAIFGGTPTTHILNLTRE 323
Db 241 TLGOLVAQLOAWYKASPEKFTYERRELSAAYIPFEGDFYVYHAAIFGGTPTTHILNLTRE 300
Qy 324 CFKGLQKKHDIEAQWHDHSHLNKYFLFNKPTKILSPYCWYDQIGLPSDIKSVKVAWQ 383
```

Db 301 CFKGLQDKKHDI EAQWHDHSLNKYFLFNKPTKILSPYCWQYQIGLPSDIKSVKVAWQ 360  
Qy 384 TKEYNLVRNV 394  
Db 361 TKEYNLVRNV 371  
RESULT 9  
US-08-378-617-12  
; Sequence 12, Application US/08378617  
; GENERAL INFORMATION:  
; APPLICANT: d'Apice, Anthony J.F.  
; APPLICANT: Pearse, Martin J.  
; APPLICANT: Robins, Allan J.  
; APPLICANT: Crawford, Robert J.  
; APPLICANT: Rathjen, Peter D.  
; TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 120 South Sixth Street, Suite 2500  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/378.617  
; FILING DATE: 26-JAN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 06868/005001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (612) 335-5070  
; TELEFAX: (612) 288-9696  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 371 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-378-617-12

Query Match 91.3%; Score 1950; DB 7; Length 371;  
Best Local Similarity 96.5%; Pred. No. 3.4e-193;  
Matches 358; Conservative 1; Mismatches 0; Indels 12; Gaps 1;  
Qy 36 MNVKGKIVLLMLIVSTVVVVFWEYVN-----RIPEVGENRWQKDMWFPSPWFKN 83  
Db 1 MNVKGKIVLLMLIVSTVVVVFWEYVNSPDGSLWYHTKIPEVGENRWQKDMWFPSPWFKN 60  
Qy 84 GTHSYQEDNVGREGKGRNGRIEPEQLWDFNPKNRPDLVTPWKAPIVWEGTYDTAL 143  
Db 61 GTHSYQEDNVGREGKGRNGRIEPEQLWDFNPKNRPDLVTPWKAPIVWEGTYDTAL 120  
Qy 144 LEKYATQKLVGLTVFAVGKYIEHYLEDLFLESADMFVWCHRVIFYVMIIDTTSRMPVH 203  
Db 121 LEKYATQKLVGLTVFAVGKYIEHYLEDLFLESADMFVWCHRVIFYVMIIDTTSRMPVH 180  
Qy 204 LNPVLSQVFEIRSEKRWQDISMMRMKTIGRHLAHIQHEVDFLFCMDVDQVQDNFQGE 263  
Db 181 LNPVLSQVFEIRSEKRWQDISMMRMKTIGRHLAHIQHEVDFLFCMDVDQVQDNFQGE 240  
Qy 264 TLGQLVAQLQAWMYKASPEKTYERRELSAAYIPFEGGDFYHAAIFGGTPTHILNLTRE 323

Db 241 TLGQLVAQLQAWMYKASPEKTYERRELSAAYIPFEGGDFYHAAIFGGTPTHILNLTRE 300  
Qy 324 CFKGLQDKKHDI EAQWHDHSLNKYFLFNKPTKILSPYCWQYQIGLPSDIKSVKVAWQ 383  
Db 301 CFKGLQDKKHDI EAQWHDHSLNKYFLFNKPTKILSPYCWQYQIGLPSDIKSVKVAWQ 360  
Qy 384 TKEYNLVRNV 394  
Db 361 TKEYNLVRNV 371  
RESULT 10  
US-08-984-900-12  
; Sequence 12, Application US/08984900  
; GENERAL INFORMATION:  
; APPLICANT: d'Apice, Anthony J.F.  
; APPLICANT: Pearse, Martin J.  
; APPLICANT: Robins, Allan J.  
; APPLICANT: Crawford, Robert J.  
; APPLICANT: Rathjen, Peter D.  
; TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 120 South Sixth Street, Suite 2500  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/984.900  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/378.617  
; FILING DATE: 26-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 06868/005001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (612) 335-5070  
; TELEFAX: (612) 288-9696  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 371 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-984-900-12

Query Match 91.3%; Score 1950; DB 13; Length 371;  
Best Local Similarity 96.5%; Pred. No. 3.4e-193;  
Matches 358; Conservative 1; Mismatches 0; Indels 12; Gaps 1;  
Qy 36 MNVKGKIVLLMLIVSTVVVVFWEYVN-----RIPEVGENRWQKDMWFPSPWFKN 83  
Db 1 MNVKGKIVLLMLIVSTVVVVFWEYVNSPDGSLWYHTKIPEVGENRWQKDMWFPSPWFKN 60  
Qy 84 GTHSYQEDNVGREGKGRNGRIEPEQLWDFNPKNRPDLVTPWKAPIVWEGTYDTAL 143  
Db 61 GTHSYQEDNVGREGKGRNGRIEPEQLWDFNPKNRPDLVTPWKAPIVWEGTYDTAL 120  
Qy 144 LEKYATQKLVGLTVFAVGKYIEHYLEDLFLESADMFVWCHRVIFYVMIIDTTSRMPVH 203



Db 121 LEKYATOKLTGVTGTVFAVGKYYIEHYLEDPLESADMYFMVGHVRVIFYVMIDDTSRMPVVH 180  
Qy 204 LNPILSLQVFEIRSEKRWQDISMMRMKTIGEHILAHIQHEVDVDFLCMDVQVQFODNFGVE 263  
Db 181 LNPILSLQVFEIRSEKRWQDISMMRMKTIGEHILAHIQHEVDVDFLCMDVQVQFODNFGVE 240  
Qy 264 TLGOLVAQLQAWMYKASPEKFTYERRELSAAYIPFGGDFYYHAAIFGGTPTTHILNLTRE 323  
Db 241 TLGOLVAQLQAWMYKASPEKFTYERRELSAAYIPFGGDFYYHAAIFGGTPTTHILNLTRE 300  
Qy 324 CFKGILODKKHDIQAQWHDHSHLNKYLEFNKPTKILSPCYCWDYQIGLPSDIKSVKVAWO 383  
Db 301 CFKGILODKKHDIQAQWHDHSHLNKYLEFNKPTKILSPCYCWDYQIGLPSDIKSVKVAWO 360  
Qy 384 TKEYNLVRNNV 394  
Db 361 TKEYNLVRNNV 371

## RESULT 11

US-09-870-759-74  
; Sequence 74, Application US/09870759  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 870759  
; CURRENT APPLICATION NUMBER: US/09/870,759  
; PRIOR APPLICATION DATE: 2002-01-14  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 166  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 74  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-870-759-74

Query Match 91.3%; Score 1950; DB 28; Length 371;  
Best Local Similarity 96.5%; Pred. No. 3.4e-193;  
Matches 358; Conservative 1; Mismatches 0; Indels 12; Gaps 1;

Qy 36 MNVKGKVVLLMLIVSTVVVFWVEYVN-----RIPEVGENRWQDMWFPSPFKN 83  
Db 1 MNVKGKVVLLMLIVSTVVVFWVEYVNSPDGSLWYHTKIPVGENRWQDMWFPSPFKN 60  
Qy 84 GTHSYQEDNVEGRKGRNGDRIEEPQLWDFNPKNRPDLVTVTPWKAPIVWEGTYDTAL 143  
Db 61 GTHSYQEDNVEGRKGRNGDRIEEPQLWDFNPKNRPDLVTVTPWKAPIVWEGTYDTAL 120  
Qy 144 LEKYATOKLTGVTGTVFAVGKYYIEHYLEDPLESADMYFMVGHVRVIFYVMIDDTSRMPVVH 203  
Db 121 LEKYATOKLTGVTGTVFAVGKYYIEHYLEDPLESADMYFMVGHVRVIFYVMIDDTSRMPVVH 180  
Qy 204 LNPILSLQVFEIRSEKRWQDISMMRMKTIGEHILAHIQHEVDVDFLCMDVQVQFODNFGVE 263  
Db 181 LNPILSLQVFEIRSEKRWQDISMMRMKTIGEHILAHIQHEVDVDFLCMDVQVQFODNFGVE 240  
Qy 264 TLGOLVAQLQAWMYKASPEKFTYERRELSAAYIPFGGDFYYHAAIFGGTPTTHILNLTRE 323  
Db 241 TLGOLVAQLQAWMYKASPEKFTYERRELSAAYIPFGGDFYYHAAIFGGTPTTHILNLTRE 300  
Qy 324 CFKGILODKKHDIQAQWHDHSHLNKYLEFNKPTKILSPCYCWDYQIGLPSDIKSVKVAWO 383  
Db 301 CFKGILODKKHDIQAQWHDHSHLNKYLEFNKPTKILSPCYCWDYQIGLPSDIKSVKVAWO 360  
Qy 384 TKEYNLVRNNV 394  
Db 361 TKEYNLVRNNV 371

## RESULT 12

US-09-994-427A-12  
; Sequence 12, Application US/09994427A  
; GENERAL INFORMATION:  
; APPLICANT: Geron Corporation  
; APPLICANT: Schiff, J. Michael  
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE VECTORS FOR TREATING CANCER  
; FILE REFERENCE: 083,002  
; CURRENT APPLICATION NUMBER: US/09/994,427A  
; CURRENT FILING DATE: 2002-02-26  
; PRIOR APPLICATION NUMBER: 60/253,395  
; PRIOR FILING DATE: 2000-11-27  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 12  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Consensus of mammalian galactosyl transferase sequences - this  
; OTHER INFORMATION: vention  
US-09-994-427A-12

Query Match 72.2%; Score 1542.5; DB 28; Length 376;  
Best Local Similarity 72.9%; Pred. No. 7.9e-151;  
Matches 274; Conservative 42; Mismatches 43; Indels 17; Gaps 2;  
Qy 36 MNVKGKVVLLMLIVSTVVVFWVEYVN-----RIPEVGENRWQDMWFPSPFKN 83  
Db 1 MNVKGKVVLLMLIVSTVVVFWVEYVNSPDGSLWYHTKIPVGENRWQDMWFPSPFKN 60  
Qy 84 GTHSYQEDNVEGRKGRNGDRIEEPQLWDFNPKNRPDLVTVTPWKAPIVWEGT 138  
Db 61 GIHNYQOEEDTDKEKGRREEQKEDDTLRLWDFNPKNRPDMVTQMKAPVWEGT 120  
Qy 139 YDTALLEKYYATOKLTGVTGTVFAVGKYYIEHYLEDPLESADMYFMVGHVRVIFYVMIDTTSR 198  
Db 121 YNKAILENYYAKOKITVGLTVFAIGRYIEHYLEEFLSANRYFMVGHVIFYVMVDVSK 180  
Qy 199 MPVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAHIQHEVDVDFLCMDVQVQFOD 258  
Db 181 APFIEGLPLKSFVFEVPEKRWQDISMMRMKTIGEHILAHIQHEVDVDFLCMDVQVQFOD 240  
Qy 259 NFGVETLGQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGGDFYYHAAIFGGTPTTHIL 318  
Db 241 HFGVETLGQSVQALQAWMYKADPDFTYERRKESAAVPEGQGFYYHAAIFGGTPTIOVL 300  
Qy 319 NLTRCEKGILODKKHDIQAQWHDHSHLNKYLEFNKPTKILSPCYCWDYQIGLPSDIKSV 378  
Db 301 NITQCEKGILODKKHDIQAQWHDHSHLNKYLEFNKPTKILSPCYCWDYHIGLPSDIKT 360  
Qy 379 KVAMOTKEYNLVRNNV 394  
Db 361 KLSWOTKEYNLVRNNV 376

## RESULT 13

US-08-213-200A-2  
; Sequence 2, Application US/08213200A  
; GENERAL INFORMATION:  
; APPLICANT: Galili, Dr. Uri  
; APPLICANT: Replik, Dr. Patricia M.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR VACCINES  
; TITLE OF INVENTION: COMPRISING ALPHA-CALACTOSYL EPITOPES  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
; STREET: Suite 1800, Two Penn Center Plaza  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19102  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE:  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/213,200A  
FILING DATE: 15-MAR-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Daniel A. Monaco  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 8760-2 US  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-213-200A-2

Query Match 72.1%; Score 1539.5; DB 6; Length 376;  
Best Local Similarity 72.6%; Pred. No. 1.6e-150;  
Matches 273; Conservative 43; Mismatches 43; Indels 17; Gaps 2;  
QY 36 MNVKGKVLILLMLIVSTVVVFWEYVN-----RIPEVGENRWQKDWFPSPFN 83  
DB 1 MNVKGKVLILLMLIVSTVVVFWEYVN-----RIPEVGENRWQKDWFPSPFN 60  
QY 84 GTHSYQE-----DNVEGRREKGRNGRIEPEQLWDWPNKRNPDVLTVPWKAPVMEGT 138  
DB 61 GIHNYQEEEDTDKEKGRREEQKEDDTTELRLWDWPNKRPVMTVQKAPVMEGT 120  
QY 139 YDTALLEKYYATQKLTVGLTVFAVGKYYIEHYLEDFLESADMYFMVGHVRVIFYVMDTTSR 198  
DB 121 YNKAILENYYAKQKITVGLTVFAIGRYIEHYLEEFVTSANRYFMVGHVIFVYVMDVDSK 180  
QY 199 MPVVHLNPLHLSQVFEIRSEKRWODISMMRMKTTIGEHILAHQHEVDVDFLCMDVDVQFOD 258  
DB 181 APFTELGLPLRSFKVFEVPEKRWQDISMMRMKTTIGEHILAHQHEVDVDFLCMDVDVQFOD 240  
QY 259 NFGVETLGLQVAQLQAWMYKASPEKFTYERRELSAAYIPFEGDFYHYHAAIFGCTPTHIL 318  
DB 241 HFGVETLGLQVAQLQAWMYKADPDFTYERRKESAAVIFPGQGFYHYHAAIFGCTPIQVL 300  
QY 319 NLTRCEKFGILQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILSPCYCDYQIGLPSDIKSV 378  
DB 301 NITQECFKGILLDKKNDIEAEWHDHSHLNKYFLFNKPSKILSPCYCDYHIGLPSDIKT 360  
QY 379 KVAQOTKEYNLVRNNV 394  
DB 361 KLSWQTKREYNLVRNNV 376

RESULT 14  
US-09-593-316-6  
SEQUENCE 6, Application US/09593316  
GENERAL INFORMATION:  
APPLICANT: Clark Mr., John  
TITLE OF INVENTION: Animal Tissue For Xenotransplantation  
FILE REFERENCE: 730/002  
CURRENT APPLICATION NUMBER: US/09/593,316  
PRIOR FILING DATE: 2000-06-13  
PRIOR APPLICATION NUMBER: 60/204,148  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 376

TYPE: PRT  
ORGANISM: Mermoset alpha 1.3-CT  
US-09-593-316-6  
Query Match 72.1%; Score 1539.5; DB 19; Length 376;  
Best Local Similarity 72.6%; Pred. No. 1.6e-150;  
Matches 273; Conservative 43; Mismatches 43; Indels 17; Gaps 2;  
QY 36 MNVKGKVLILLMLIVSTVVVFWEYVN-----RIPEVGENRWQKDWFPSPFN 83  
DB 1 MNVKGKVLILLMLIVSTVVVFWEYVN-----RIPEVGENRWQKDWFPSPFN 60  
QY 84 GTHSYQE-----DNVEGRREKGRNGRIEPEQLWDWPNKRNPDVLTVPWKAPVMEGT 138  
DB 61 GIHNYQEEEDTDKEKGRREEQKEDDTTELRLWDWPNKRPVMTVQKAPVMEGT 120  
QY 139 YDTALLEKYYATQKLTVGLTVFAVGKYYIEHYLEDFLESADMYFMVGHVRVIFYVMDTTSR 198  
DB 121 YNKAILENYYAKQKITVGLTVFAIGRYIEHYLEEFVTSANRYFMVGHVIFVYVMDVDSK 180  
QY 199 MPVVHLNPLHLSQVFEIRSEKRWODISMMRMKTTIGEHILAHQHEVDVDFLCMDVDVQFOD 258  
DB 181 APFTELGLPLRSFKVFEVPEKRWQDISMMRMKTTIGEHILAHQHEVDVDFLCMDVDVQFOD 240  
QY 259 NFGVETLGLQVAQLQAWMYKASPEKFTYERRELSAAYIPFEGDFYHYHAAIFGCTPTHIL 318  
DB 241 HFGVETLGLQVAQLQAWMYKADPDFTYERRKESAAVIFPGQGFYHYHAAIFGCTPIQVL 300  
QY 319 NLTRCEKFGILQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILSPCYCDYQIGLPSDIKSV 378  
DB 301 NITQECFKGILLDKKNDIEAEWHDHSHLNKYFLFNKPSKILSPCYCDYHIGLPSDIKT 360  
QY 379 KVAQOTKEYNLVRNNV 394  
DB 361 KLSWQTKREYNLVRNNV 376  
RESULT 15  
US-09-173-270-2  
SEQUENCE 2, Application US/09173270  
GENERAL INFORMATION:  
APPLICANT: GALILI, URI  
TITLE OF INVENTION: REPIK, PATRICIA M.  
COMPOSING ALPHA-GALACTOSYL EPITOPES  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
STREET: Suite 1800, Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/173,270  
FILING DATE: 23-Mar-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/704,548  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 8760-2 CII  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-173-270-2

Query Match	72.1%	Score 1539.5	DB 27	Length 376
Best Local Similarity	72.6%	Pred. No. 1.6e-150		
Matches 273	Conservative 43	Mismatches 43	Indels 17	Gaps 2

  

QY	36	MNVKGVILMLIVSVVVVFWYVN-----RIPEYGENRWQKDMWFPSWFKN	83
DB	1	MNVKGVILMLVSVIVVFWYINSPGSLVIYHNSKNPEYDDSSAQKDMWFPGWNN	60
QY	84	GTHSYOE-----DNVGRREKGRNGRIEPPQLMDWPNKRNPDVLTVPWKAPIVWEGT	138
DB	61	GIHNYQOEEDTDKERGEEQKEDDTTLRLMDWPNKKRPEVMTVTQWKAPVWEGT	120
QY	139	YDTALLEKYVATOKLTAVGLTVFAVGKYIEHYLEDFLESADMYFMVGHVIFVYVMIIDTSR	198
DB	121	YNKAILENYAKOKITVGLTVFAIGRYIEHYLEEFVTSANRYFMVGHKIFYVMVDDVSK	180
QY	199	MPVYHLNPLHSLQVFEIRSEKRWQDISMMRMKTTIGEHILAHIQHEVDFLCMDVDQVQD	258
DB	181	APFIELGPLRSFKVFEKPKRWQDISMMRMKTTIGEHILAHIQHEVDFLCMDVDQVQD	240
QY	259	NFGVETLGQLVAQLQAWMYKASPEKTYERRELSAAYIPFGEGDFYHAAIFGCTPTHIL	318
DB	241	HFGVETLGQSVQALQAWMYKADPDDETYERRKESAAAYIPFGQDFYHAAIFGCTPIQVL	300
QY	319	NLTRECFCGTLQDKKHIDIEAQWHDSEHLNKYFLFNKPTKILSPCYCWDYQIGLPSDIKS	378
DB	301	NITQECFCGILLDDKNDIEAWEHDEHLNKYFLLNKPSKILSPCYCWDYHIGLPSDIKT	360
QY	379	KVARQTKYENLVRNV	394
DB	361	KLSMQTKYENLVRKNV	376

Search completed: May 10, 2002, 11:09:25
Job time: 208 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:06:22 ; Search time 11.66 Seconds  
(without alignments)  
720.742 Million cell updates/sec

Title: US-09-863-475A-4  
Perfect score: 2136  
Sequence: 1 MITMLQDLHVKNISMSRSK.....IKSVKVAWQTKYNLVRNV 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 92415 seqs, 21329587 residues

Total number of hits satisfying chosen parameters: 92415

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	229	10.7	106	6	US-10-115-123-419 Sequence 419, App
2	175	8.2	90	6	US-10-115-123-190 Sequence 190, App
3	175	8.2	90	6	US-10-115-123-240 Sequence 240, App
4	175	8.2	90	6	US-10-115-123-417 Sequence 417, App
5	97	4.5	432	5	US-09-602-777A-140 Sequence 140, App
6	86	4.0	887	1	PCT-US02-09288-20 Sequence 20, Appl
7	85.5	4.0	368	1	PCT-US02-09944-432 Sequence 432, App
8	85.5	4.0	662	5	US-09-540-209B-8400 Sequence 8400, App
9	84	3.9	746	6	US-10-041-007-39 Sequence 39, Appl
10	84	3.9	795	6	US-10-041-007-37 Sequence 37, Appl
11	84	3.9	814	6	US-10-041-007-35 Sequence 35, Appl
12	84	3.9	873	6	US-10-041-007-2 Sequence 2, Appl
13	84	3.9	873	6	US-10-041-007-33 Sequence 33, Appl
14	84	3.9	873	6	US-10-041-018-398 Sequence 398, App
15	82.5	3.9	696	5	US-09-540-209B-9947 Sequence 9947, App
16	79.5	3.7	379	5	US-09-540-209B-6269 Sequence 6269, App
17	78.5	3.7	417	5	US-09-540-209B-8036 Sequence 8036, App
18	77.5	3.6	2214	1	PCT-US02-07826-300 Sequence 300, App
19	77.5	3.6	2214	1	PCT-US02-09671-810 Sequence 810, App
20	77.5	3.6	2214	1	PCT-US02-09671-811 Sequence 811, App
21	77.5	3.6	2214	1	PCT-US02-09671-812 Sequence 812, App
22	77.5	3.6	2214	1	PCT-US02-09671-813 Sequence 813, App
23	77.5	3.6	2214	1	PCT-US02-09671-814 Sequence 814, App
24	77.5	3.6	2214	6	US-10-097-340-300 Sequence 300, App
25	76	3.6	249	5	US-09-540-209B-5225 Sequence 5225, App
26	75.5	3.5	690	5	US-09-540-209B-6647 Sequence 6647, App

ALIGNMENTS

RESULT 1

US-10-115-123-419  
; Sequence 419, Application US/10115123  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30AP1D2  
; CURRENT APPLICATION NUMBER: US/10/115,123  
; PRIOR FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 419  
; LENGTH: 106  
; TYPE: PRT  
; ORGANISM: Homo sapiens

27	75.5	3.5	1722	1	PCT-US02-09671-666	Sequence 666, App
28	75.5	3.5	1722	1	PCT-US02-09671-667	Sequence 667, App
29	75.5	3.5	1722	1	PCT-US02-09671-670	Sequence 670, App
30	75.5	3.5	1722	1	PCT-US02-09671-671	Sequence 671, App
31	75.5	3.5	1722	1	PCT-US02-09671-672	Sequence 672, App
32	75.5	3.5	1722	1	PCT-US02-09671-673	Sequence 673, App
33	75.5	3.5	1743	1	PCT-US02-09671-668	Sequence 668, App
34	75	3.5	224	5	US-09-540-209B-6438	Sequence 6438, App
35	75	3.5	471	5	US-09-573-655B-945	Sequence 945, App
36	75	3.5	1428	5	US-09-573-655B-1944	Sequence 1944, App
37	74.5	3.5	392	5	US-09-540-209B-9064	Sequence 9064, App
38	74.5	3.5	406	5	US-09-591-521-6	Sequence 6, Appl
39	74.5	3.5	837	5	US-09-540-209B-8217	Sequence 8217, App
40	74.5	3.5	1307	5	US-09-573-655B-1888	Sequence 1888, App
41	74	3.5	468	5	US-09-540-209B-8955	Sequence 8955, App
42	74	3.5	469	1	PCT-US02-10824-164	Sequence 164, App
43	74	3.5	469	1	PCT-US02-10421-2912	Sequence 2912, App
44	74	3.5	469	6	US-10-112-699-2912	Sequence 2912, App
45	74	3.5	485	5	US-09-540-209B-5981	Sequence 5981, App

Query Match 10.7%; Score 229; DB 6; Length 106;  
Best Local Similarity 52.4%; Pred. No. 1.3e-14;  
Matches 43; Conservative 12; Mismatches 15; Indels 12; Gaps 1;

Qy 32 OEKTMNVKGVILLMLIVSTVVVWFVYVN-----RIPEVGENRWQKDMPPS 79  
||| ||||| ||:||||:||||:|  
Db 3 QEKIRNVKGVILSMLVSTVIVWFVFINSTEGSFLWVHSKNPEVDSSAQKGNWFLS 62

Qy 80 WFKNGTHSYQEDNVGREGKR 101  
|| ||||: : ||||  
Db 63 WFNNGIHHYQOGEEDIDKEGR 84

RESULT 2

US-10-115-123-190  
; Sequence 190, Application US/10115123  
; GENERAL INFORMATION:

APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30AP1D2  
; CURRENT APPLICATION NUMBER: US/10/115.123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 190  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-115-123-190

Query Match 8.2% Score 175; DB 6; Length 90;  
Best Local Similarity 45.6% Pred. No. 1.2e-09;  
Matches 31; Conservative 12; Mismatches 13; Indels 12; Gaps 1;  
  
Qy 46 MLIVSTVVVFEWYVN-----RIPEVGENRQKDWFPSPFKNGTHSYQEDNV 93  
||:||||:||||:| : ||| : || ||| ||| ||| :|||  
Db 1 MLVSTVIIVFEWFINSTEGSFLWIYHSKNPEVDDSSAQKGMWFLSWFNNGIHNYYQOGE 60  
: ||| : || ||| ||| ||| ||| :|||  
  
Qy 94 EGRREKGR 101  
: : |||||  
Db 61 DIDREKGR 68

RESULT 3  
US-10-115-123-240  
; Sequence 240, Application US/10115123  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30AP1D2  
; CURRENT APPLICATION NUMBER: US/10/115.123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 240  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-115-123-240

Query Match 8.2% Score 175; DB 6; Length 90;  
Best Local Similarity 45.6% Pred. No. 1.2e-09;  
Matches 31; Conservative 12; Mismatches 13; Indels 12; Gaps 1;  
  
Qy 46 MLIVSTVVVFEWYVN-----RIPEVGENRQKDWFPSPFKNGTHSYQEDNV 93  
||:||||:||||:| : ||| : || ||| ||| ||| :|||  
Db 1 MLVSTVIIVFEWFINSTEGSFLWIYHSKNPEVDDSSAQKGMWFLSWFNNGIHNYYQOGE 60  
: ||| : || ||| ||| ||| ||| :|||  
  
Qy 94 EGRREKGR 101  
: : |||||  
Db 61 DIDREKGR 68  
  
RESULT 4  
US-10-115-123-417  
; Sequence 417, Application US/10115123  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30AP1D2  
; CURRENT APPLICATION NUMBER: US/10/115.123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 417  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-115-123-417

Query Match 8.2% Score 175; DB 6; Length 90;  
Best Local Similarity 45.6% Pred. No. 1.2e-09;  
Matches 31; Conservative 12; Mismatches 13; Indels 12; Gaps 1;  
  
Qy 46 MLIVSTVVVFEWYVN-----RIPEVGENRQKDWFPSPFKNGTHSYQEDNV 93  
||:||||:||||:| : ||| : || ||| ||| ||| :|||  
Db 1 MLVSTVIIVFEWFINSTEGSFLWIYHSKNPEVDDSSAQKGMWFLSWFNNGIHNYYQOGE 60  
: ||| : || ||| ||| ||| ||| :|||  
  
Qy 94 EGRREKGR 101  
: : |||||  
Db 61 DIDREKGR 68  
  
RESULT 5  
US-09-602-777A-140  
; Sequence 140, Application US/09602777A  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Markus  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schroder, Hartwig  
; APPLICANT: Zelder, Oskar  
; APPLICANT: Habermann, Gregor  
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
; FILE REFERENCE: BGI-128CP  
; CURRENT APPLICATION NUMBER: US/09/602,777A  
; CURRENT FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/141031

PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: DE 19931636.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19932125.6  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932126.4  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932127.2  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932128.0  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932129.9  
PRIOR FILING DATE: 1999-07-19  
PRIOR APPLICATION NUMBER: DE 19932226.0  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932920.6  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932922.2  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932924.9  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932928.1  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932930.3  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932933.8  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932935.4  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932973.7  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933002.6  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933003.4  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933005.0  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933006.9  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19941378.9  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941379.7  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941390.8  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941391.6  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19942088.2  
PRIOR FILING DATE: 1999-09-03  
NUMBER OF SEQ ID NOS: 442  
SEQ ID NO 140  
LENGTH: 432  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-602-777A-140

Query Match 4.5%; Score 97; DB 5; Length 432;  
Best Local Similarity 18.5%; Pred. No. 0.19;  
Matches 27; Conservative 27; Mismatches 53; Indels 38; Gaps 3;  
Qy 12 KISMRSKSETSLPSSRSGSGOEKTMNVKGVILLMLIVS-----TVV 53  
Db 209 KILQAEGEKHAISILNABEAOAMILTRAEGERAARYLQAGARAIQVNAAIKSAKLTP 268  
Qy 54 VVFHEYVNRIPVEGCENRWQDNWFPSPFNK-----GTHSQEDNVEGRRE 98  
Db 269 VLAYOYLEKLPKTAIEAGNASKMWIIPSPDSLEGFAGQFQAKDAEGVFRPEPNTVDETR 328  
Qy 99 KGRNGDRIEPQLDWDFNPKNRDPV 123  
Db 329 DIANADNVE-----DNFSTESDPEI 348

RESULT 6  
PCT-US02-09288-20  
Sequence 20, Application PC/TUS0209288  
GENERAL INFORMATION:  
APPLICANT: INCYTE GENOMICS, INC.  
APPLICANT: HAFALIA, April J.A.  
APPLICANT: TANG, Y. Tom  
APPLICANT: YUE, Henry  
APPLICANT: KHAN, Farrah A.  
APPLICANT: ISON, Craig H.  
APPLICANT: BAUGHN, Mariah R.  
APPLICANT: WARREN, Bridget A.  
APPLICANT: DUGGAN, Brendan M.  
APPLICANT: THANGAVELU, Kavitha  
APPLICANT: HONCHELL, Cynthia D.  
APPLICANT: AZIMZAI, Yalda  
APPLICANT: ELLIOTT, Vicki S.  
APPLICANT: BURFORD, Neil  
APPLICANT: DING, Li  
APPLICANT: YUE, Huibin  
APPLICANT: BECHA, Shanya  
APPLICANT: EMERLING, Brooke M.  
APPLICANT: RICHARDSON, Thomas W.  
APPLICANT: LEE, Soo Yeun  
APPLICANT: BANDMAN, Olga  
APPLICANT: LAL, Preeti G.  
APPLICANT: LEE, Sally  
APPLICANT: GIETZEN, Kimberly J.  
APPLICANT: WALIA, Narinder K.  
APPLICANT: GRIFFIN, Jennifer A.  
APPLICANT: LEE, Ernestine A.  
APPLICANT: SWARNAKAR, Anita  
APPLICANT: RING, Huijun Z.  
APPLICANT: JONES, Karen Anne  
TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS  
FILE REFERENCE: PF-0918 PCT  
CURRENT APPLICATION NUMBER: PCT/US02/09288  
CURRENT FILING DATE: 2002-03-25  
PRIOR APPLICATION NUMBER: 60/280,508; 60/281,323; 60/283,769; 60/288,609; 60/290,560/291,870; 60/294,451  
PRIOR FILING DATE: 2001-03-29; 2001-04-03; 2001-04-13; 2001-05-04; 2001-05-10;  
2001-05-18; 2001-05-29  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: PERL Program  
SEQ ID NO 20  
LENGTH: 887  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No: 1990526CD1  
PCT-US02-09288-20

Query Match 4.0%; Score 86; DB 1; Length 887;  
Best Local Similarity 19.0%; Pred. No. 5.3;  
Matches 82; Conservative 58; Mismatches 138; Indels 154; Gaps 21;  
Qy 32 OEKTMNVKGVILLMLIVSVVVVFVYVNRIPVEGVENRWQKQ----- 74  
Db 83 OEKISELQN-----YQRINHFPGMGE-ICKKDFLARNMTKMKSRPLDY 125  
Qy 75 -----WVFPSWFKNGTHSQEDNVEGRREK-----RNGDR----- 105  
Db 126 TFVPRTWIFPAEYTO-FQNVYKELKKRKKQKFTFKPANGAMGHCISLRNGDKLPQSDH 184  
Qy 106 -----IEEPQMDWFPNKNRPDVLTVT--PWKAPIVWEGTYDTALLEKYATO-----K 152  
Db 185 LIVQEIYEKPFMEGYKFDLRIYILVTSQDPLKIFLYHDGLVRMG-TEKYIPPNESNLTO 243  
Qy 153 LTVGLTVFVANGKYIEHYLEDFESADMYFVWGHVIFVYVMDITSRPVPVHLNPLHSLOV 212

```
Db 244 LYMHLTNTSVKHNHFRDRTENKG-----SKRSIKWFTPE--LQA 283
Qy 213 FEIRSEKRWODISMRRMKTIGEHLA--HIQHEVDLFCMDVDQVQDNFGVETLG---Q 267
Db 284 NOHDVAKFMSDISLVKTL---IVAEPVHLA---YRRCRQPPGSGESVCFVLGFDIL 338
Qy 268 LVAQLQAWYK-----ASPEKTYE-----RRELS-----AAY 295
Db 339 LDRKLKPLLEINRAPSGTDQKIDYDVKRGVLLNALKLLNIRTSDKRRNLAKAKAEQR 398
Qy 296 IPFGEEDFYHAATFGCT-----PTHILNLTRECFKGLQDKKHDIEAQWHDHSHLNKYF 350
Db 399 RLYGONSI---KRLPGSSDWEQHQHLERKEELKERLAQVRKQISREEHENRMHNGY- 454
Qy 351 LFNKPTKILSPE 362
Db 455 -----RRIYPPE 461

RESULT 7
PCT-US02-09944-432
; Sequence 432, Application PC/TUS0209944
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DAFFO, Abel
; APPLICANT: JONES, Anissa L.
; APPLICANT: TRAN, Alanna-Phung B.
; APPLICANT: DAHL, Christopher R.
; APPLICANT: GIETZEN, Darryl
; APPLICANT: CHINN, Joyce
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: TUASON, Olivia
; APPLICANT: YAP, Pierre E.
; APPLICANT: AMSHEY, Stefan R.
; APPLICANT: DAUGHERTY, Sean C.
; APPLICANT: DAM, Tam C.
; APPLICANT: LIU, Tommy F.
; APPLICANT: KLEEFELD, Yael
; APPLICANT: NGUYEN, Duy-Viet An
; APPLICANT: GERSTIN, JR., Edward H.
; APPLICANT: PERALTA, Careyna H.
; APPLICANT: DAVID, Marie H.
; APPLICANT: LEWIS, Samantha A.
; APPLICANT: CHEN, Alice J.
; APPLICANT: PANZER, Scott R.
; APPLICANT: HARRIS, Bernard
; APPLICANT: FLORES, Vincent
; APPLICANT: MARWAHA, Rakesh
; APPLICANT: LO, Audrey
; APPLICANT: LAN, Ruth Y.
; APPLICANT: URASHKA, Michael E.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1231 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/09944
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 60/280,067; 60/279,619; 60/280,068; 60/291,280; 60/291,849;
60/291,829; 60/299,428; 60/300,001; 60/299,776
; PRIOR FILING DATE: 2001-03-29; 2001-03-28; 2001-03-29; 2001-05-16; 2001-05-17;
2001-05-17; 2001-06-19; 2001-06-20; 2001-06-20
; NUMBER OF SEQ ID NOS: 792
; SOFTWARE: PERL Program
; SEQ ID NO 432
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: LG:118834.9.orf1:2001MAR30
PCT-US02-09944-432
```

```
Query Match 4.0%; Score 85.5; DB 1; Length 368;
Best Local Similarity 23.6%; Pred. No. 1.8;
Matches 62; Conservative 40; Mismatches 84; Indels 77; Gaps 18;

Qy 152 KLTVCLTVFAVCKYIEHYLEDLFESADMYFVGHVRVI-FYVMIDDTSRMPVHLNPLHSL 210
Db 32 ELMIHLAVVACGNRLLEETLV-MLKSAVLF---SHRKIQHFIFTEDS-----LKPEFDK 80
Qy 211 QV-----FEIRSEKRWODI-----SMRMKKTIGEHLAHIQHEVDL 247
Db 81 QLQWPDSTYTKKFEHRIYPIITFSVGNPOEKKLFPKCAQRL-----FLPVILKDVDSL 134
Qy 248 FCMDDVQVQDNFGVETLGQLV-----AQLQAWYKASPEKTYERRELSAAYIPFGECD 302
Db 135 LYVDTDLVFLR--PVDDIWKLLRLFNSTQLAA---MAPE---HEIPKI-CWYSRFARHP 184
Qy 303 FYHAAIFGGTPTHILNLTTR---ECFKGL-----OD-----KKHIDIEAOWHDESHL 346
Db 185 FYSAGVNSG--VMLMNLTRISTQFKNSMPTGLAWEDMLYPLYQKYKNAITMGDQDL 242
Qy 347 NKYFLFNKPTKILSPEYC-WDYQ 368
Db 243 NIIFYNPECLYVFP--CQWYR 263

RESULT 8
US-09-540-209B-8400
; Sequence 8400, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540.209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 8400
; LENGTH: 662
; TYPE: PRT
; ORGANISM: B.fragilis
US-09-540-209B-8400

Query Match 4.0%; Score 85.5; DB 5; Length 662;
Best Local Similarity 20.8%; Pred. No. 4;
Matches 59; Conservative 36; Mismatches 86; Indels 103; Gaps 16;

Qy 164 KYIEHYLEDLFESAD-----MYFMVGHVRVIFYVMIDDTSRMPVHLNPLHSLQVFEIRS 217
Db 54 KEIEKALDFFINQKDSLKIRSIFFLVGN-----MADKYSLTTPANEQDPFHSLNNHIK 107
Qy 218 EKRWDISMRRM-----KT-----IGEHILAHIOHEVDL----- 247
Db 108 EKEAWDPCKSRGLGNALDSVYKTCDDPPRPKIVRDIEVTGNFLINNVEAEIKIHWRTK 167
Qy 248 -----FCMDVDVQVQDNFGVETLGQVLAQLQAWYKASPEKTY-----ERRELSAA 294
Db 168 TECSFDDFC---EVLTPYRIGNESLSA-----WREQACOKESYLLDSDPLDLTKA 216
Qy 295 YIPFGEEDFYHAAT--FCGTPTHILNLTRECFKGLQDKKHDIEAOWHDESHLNKYFLF 352
Db 217 IVQV--SGIYYNAGMSKYFPFPT-----FSEL--DQLH-----WGSCHDLAAAYLTF 258
Qy 353 NKPTKILSPEYCDYQIGLPSDIKSVKVAWQTK-----EYNLVRN 392
Db 259 SLRA-----IGIPSTI-DVVPWANRGGGHWNVVMN 289

RESULT 9
US-10-041-007-39
; Sequence 39, Application US/10041007
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seichi P.T.
```



APPLICANT: Schepmann, Hala G  
TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
FILE REFERENCE: P02081US1  
CURRENT APPLICATION NUMBER: US/10/041,007  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: US 60/259,881  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 39  
LENGTH: 746  
TYPE: PRT  
ORGANISM: Ginkgo biloba  
US-10-041-007-39

Query Match 3.9%; Score 84; DB 6; Length 746;  
Best Local Similarity 17.6%; Pred. No. 6.5; Mismatches 116; Indels 150; Gaps 17;  
Matches 68; Conservative 53;  
QY 109 PQLWDFNPKNRDPDLTVTPWKAPIVWEGTYDTALLEKYVATOKLTVGLTVFAVG----- 163  
DB 35 PQLWILNNQLPD-----GSMGECIFL-AYDRVL-----NTLACLLTLKINWKGDIQVQ 84  
QY 164 ---KYIEHYLEDLESADMYFMVGHVRVIFYVMIDDT-----SR 198  
DB 85 KGVEFVRKHMEEMKDEADNHRPSGFVFPAMLDKSLGLDLPYHLPFISQIHQKROKK 144  
QY 199 MPVVHLNPLHSLQVFEIRSEK-----RWODISMNR----- 228  
DB 145 LQKIPLNVLNHNQVALLYSLGLEQDVVDVDMQEIITNLQSRDGSFSSPASTACVFMHTQNK 204  
QY 229 -----MKTIGEHILAHIQHE-----VDFLCMDVDQVQDNFQVETLGLQV----- 269  
DB 205 CLHFLNVLKSGDGYVPCHPYPLDLFERLWAVDTVERLGIDRYPKKEI-KESLDVYVRYWD 263  
QY 270 AQLOANWYKASP-----EFTYERRELSAAYIPFEGGDF 303  
DB 264 AERGVGWARGNPIPDVDDTAMGLRLILRHGYNVSSDVLENFRDEK-----GDF 311  
QY 304 YYHAAIFGGTPTHTLNLTNR-----ECFKGILQDKKHDIKHAOWHDESHLN-----KYFLFNKPT 356  
DB 312 FCAGQTQIGVTGNLNLRYCSQVCFPG-----EKIMEAKFTTTNHLQNALAKNAFADK-- 365  
QY 357 KILSPCYWDYQICLPSDIK-SVKVAV 382  
DB 366 -----WAVKKDLPCGEVEVAIKPW 384

RESULT 10  
US-10-041-007-37  
Sequence 37, Application US/10041007  
GENERAL INFORMATION:  
APPLICANT: Matsuda, Seichi P.T.  
APPLICANT: Schepmann, Hala G  
TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
FILE REFERENCE: P02081US1  
CURRENT APPLICATION NUMBER: US/10/041,007  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: US 60/259,881  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 37  
LENGTH: 795  
TYPE: PRT  
ORGANISM: Ginkgo biloba  
US-10-041-007-37

Query Match 3.9%; Score 84; DB 6; Length 795;  
Best Local Similarity 17.6%; Pred. No. 7.1;  
Matches 68; Conservative 53; Mismatches 116; Indels 150; Gaps 17;

QY 109 PQLWDFNPKNRDPDLTVTPWKAPIVWEGTYDTALLEKYVATOKLTVGLTVFAVG----- 163  
DB 84 PQLWILNNQLPD-----GSMGECIFL-AYDRVL-----NTLACLLTLKINWKGDIQVQ 133  
QY 164 ---KYIEHYLEDLESADMYFMVGHVRVIFYVMIDDT-----SR 198  
DB 134 KGVEFVRKHMEEMKDEADNHRPSGFVFPAMLDKSLGLDLPYHLPFISQIHQKROKK 193  
QY 199 MPVVHLNPLHSLQVFEIRSEK-----RWODISMNR----- 228  
DB 194 LQKIPLNVLNHNQVALLYSLGLEQDVVDVDMQEIITNLQSRDGSFSSPASTACVFMHTQNK 253  
QY 229 -----MKTIGEHILAHIQHE-----VDFLCMDVDQVQDNFQVETLGLQV----- 269  
DB 254 CLHFLNVLKSGDGYVPCHPYPLDLFERLWAVDTVERLGIDRYPKKEI-KESLDVYVRYWD 312  
QY 270 AQLOANWYKASP-----EFTYERRELSAAYIPFEGGDF 303  
DB 313 AERGVGWARGNPIPDVDDTAMGLRLILRHGYNVSSDVLENFRDEK-----GDF 360  
QY 304 YYHAAIFGGTPTHTLNLTNR-----ECFKGILQDKKHDIKHAOWHDESHLN-----KYFLFNKPT 356  
DB 361 FCAGQTQIGVTGNLNLRYCSQVCFPG-----EKIMEAKFTTTNHLQNALAKNAFADK-- 414  
QY 357 KILSPCYWDYQICLPSDIK-SVKVAV 382  
DB 415 -----WAVKKDLPCGEVEVAIKPW 433

RESULT 11  
US-10-041-007-35  
Sequence 35, Application US/10041007  
GENERAL INFORMATION:  
APPLICANT: Matsuda, Seichi P.T.  
APPLICANT: Schepmann, Hala G  
TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
FILE REFERENCE: P02081US1  
CURRENT APPLICATION NUMBER: US/10/041,007  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: US 60/259,881  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 35  
LENGTH: 814  
TYPE: PRT  
ORGANISM: Ginkgo biloba  
US-10-041-007-35

Query Match 3.9%; Score 84; DB 6; Length 814;  
Best Local Similarity 17.6%; Pred. No. 7.3;  
Matches 68; Conservative 53; Mismatches 116; Indels 150; Gaps 17;

QY 109 PQLWDFNPKNRDPDLTVTPWKAPIVWEGTYDTALLEKYVATOKLTVGLTVFAVG----- 163  
DB 103 PQLWILNNQLPD-----GSMGECIFL-AYDRVL-----NTLACLLTLKINWKGDIQVQ 152  
QY 164 ---KYIEHYLEDLESADMYFMVGHVRVIFYVMIDDT-----SR 198  
DB 153 KGVEFVRKHMEEMKDEADNHRPSGFVFPAMLDKSLGLDLPYHLPFISQIHQKROKK 212  
QY 199 MPVVHLNPLHSLQVFEIRSEK-----RWODISMNR----- 228  
DB 213 LQKIPLNVLNHNQVALLYSLGLEQDVVDVDMQEIITNLQSRDGSFSSPASTACVFMHTQNK 272  
QY 229 -----MKTIGEHILAHIQHE-----VDFLCMDVDQVQDNFQVETLGLQV----- 269  
DB 273 CLHFLNVLKSGDGYVPCHPYPLDLFERLWAVDTVERLGIDRYPKKEI-KESLDVYVRYWD 331  
QY 270 AQLOANWYKASP-----EFTYERRELSAAYIPFEGGDF 303



```
Db 162 POTLQWLTNNQLPD-----GSWGECIEFL-AYDRVL-----NTLACLTLTKIWNKGDIOVQ 211
QY 164 ---KYIEHYLEDFLESADMYFMVGHRYFVYMIDDT-----SR 198
Db 212 KGVEFVRKHMEKDEADNHRPSCFEVVPAMLDEAKSLGLDLPYHLPFISQIHQKQKK 271
QY 199 MPVVHLNPLHSLQVFEIRSEK-----RWQDISMMR-----228
Db 272 LQKIPLNLVNLHHQTALYSLEGLQDVVDQWQEIITNLQSRGDSFLSSPASTACVFEMHTQNK 331
QY 229 -----MKTIGEHILAHIQHE-----VDFLCMDVDQVQDFQDFVETLGLV----269
Db 332 CLHFLNVLKFGDYVPCHPYPLDLFERLMAVDVVERLGIDRYFAKEI-KESLOYVRYWD 390
QY 270 AOLQAMMYKASP-----EKFTYERRELSAAYIPFGEQDF 303
Db 391 AERGVGWARCNPIDVDVDTAMGLRLRLHGYNVSSDVLENFROEK-----GDF 438
QY 304 YYHAAIFGGTPTTHLNLTR---ECFKILQDKKHIDIAQWHDHSHLN-----KYFLFNKPT 356
Db 439 FCFAGOTQIGVTDNLNLYRCQVCFPG-----EKIMEEAKFTTTHLQNALAKNNAFDK--492
QY 357 KILSPEYCDYOIGLPSDIK-SVKVAW 382
Db 493 -----WAVKKDLPGVEVETAIKYPW 511

RESULT 15
US-09-540-209B-9947
; Sequence 9947, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540, 209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 9947
; LENGTH: 696
; TYPE: PRT
; ORGANISM: B.fragilis
US-09-540-209B-9947
```

```
Query Match 3.9%; Score 82.5; DB 5; Length 696;
Best Local Similarity 21.5%; Pred. No. 8.2;
Matches 34; Conservative 17; Mismatches 48; Indels 59; Gaps 7;

QY 40 GKVILLMLI---VSTVVVVFWEYVYVNRIPVEGEN-----69
Db 411 GKQOLIELIRQVHNHPSICFWGLFNLKEVGDNPVEYVKELNLAKQEDPTRPTTSASNQ 470
QY 70 -----RWQK-DWVFPSS-----WPKNGTHSYQEDNVGCRREKGRNG-----DR 105
Db 471 DGNLNFITENIANNRYDCWGSTPKTLATFLDRTHKKHPELRIGISEYGAGASYHQQDS 530
QY 106 IEEPOLWDFNPKNRDPVLVTPWK-----APIVWEGTY 139
Db 531 LKQPSASGWHHPENWQTYYYHWNKIIAERPFVW-GTF 567
```

Search completed: May 10, 2002, 11:09:44  
Job time: 202 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 18:40:43 ; Search time 129.59 Seconds  
(without alignments)  
5418.324 Million cell updates/sec

Title: US-09-863-475A-3  
Perfect score: 1500  
Sequence: 1 CTTCCCTTGTAGACTTTC.....CTTGACACTATTCTTAACCA 1500

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 424068 seqs, 234053524 residues

Total number of hits satisfying chosen parameters: 848136

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_New.\*

- 1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	98.8	6.6	1150	6	US-10-115-123-69
2	85.4	5.7	1193	6	US-10-115-123-119
3	84.2	5.6	305	5	US-09-540-2108-12481
4	37.8	2.5	3186	6	US-10-102-524-1730
5	36.8	2.5	8809	5	US-09-053-375B-710
6	36.2	2.4	790	6	US-10-123-155-204
7	36.2	2.4	837	5	US-09-647-041C-7
8	36.2	2.4	3328	5	US-09-053-375B-494
9	36	2.4	1049	6	US-10-123-155-358
10	35.6	2.4	658	6	US-10-102-806-97
11	35.4	2.4	39827	6	US-10-041-018-399
12	35.2	2.3	321	5	US-09-736-968A-80
13	35.2	2.3	547	5	US-09-312-283B-14
14	35.2	2.3	6372	5	US-09-736-968A-1
15	34.6	2.3	605	6	US-10-123-155-160
16	34.2	2.3	933	6	US-10-040-647-29
17	34.2	2.3	1635	1	PCT-US02-09785-73
18	34	2.3	240	5	US-09-975-254-212
19	34	2.3	243	5	US-09-540-2108-6095
20	34	2.3	634	6	US-10-108-580-8
21	33.6	2.2	1726	1	PCT-US02-08992-13
22	33.4	2.2	1184	6	PCT-US02-11682-3
23	33.2	2.2	4500	1	PCT-US02-11682-3
24	33.2	2.2	4500	6	US-10-122-087-3
25	33.2	2.2	6568	1	PCT-US02-11682-1
26	33.2	2.2	6568	6	US-10-122-067-1

27	33.2	2.2	7297	6	US-10-105-299-10267	Sequence 10267, A
28	33.2	2.2	8420	6	US-10-105-299-10266	Sequence 10266, A
29	33	2.2	594	6	US-10-123-155-10	Sequence 10, Appl
30	33	2.2	2288	6	US-10-105-299-2436	Sequence 2436, Ap
31	33	2.2	3406	6	US-10-105-299-13185	Sequence 13185, A
32	33	2.2	3406	6	US-10-105-299-13186	Sequence 13186, A
33	32.4	2.2	897	5	US-09-540-209B-4748	Sequence 4748, Ap
34	32.2	2.1	1659	6	US-10-128-714-2541	Sequence 2541, Ap
35	32.2	2.1	1659	6	US-10-128-714-7541	Sequence 7541, Ap
36	32.2	2.1	2001	6	US-10-128-714-1541	Sequence 1541, Ap
37	32.2	2.1	2001	6	US-10-128-714-6541	Sequence 6541, Ap
38	32.2	2.1	3930	1	PCT-US02-09288-39	Sequence 39, Appl
39	32.2	2.1	4001	6	US-10-128-714-541	Sequence 541, App
40	32.2	2.1	4001	6	US-10-128-714-5541	Sequence 5541, Ap
41	32	2.1	1580	5	US-09-053-375B-13	Sequence 13, Appl
42	32	2.1	3939	6	US-10-006-063A-226	Sequence 226, App
43	32	2.1	3939	6	US-10-006-117A-226	Sequence 226, App
44	32	2.1	3939	6	US-10-006-130A-226	Sequence 226, App
45	32	2.1	3939	6	US-10-006-172A-226	Sequence 226, App

## ALIGNMENTS

### RESULT 1

US-10-115-123-69  
; Sequence 69, Application US/10115123  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30AP1D2  
; CURRENT APPLICATION NUMBER: US/10/115,123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1998-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 69  
; LENGTH: 1150  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-115-123-69

Query Match 6.6%; Score 98.8; DB 6; Length 1150;

Best Local Similarity 65.5%; Pred No. 4.8e-19;

Matches 203; Conservative 0; Mismatches 57; Indels 50; Gaps 2;

QY	370	CAGGAGAAATAATGAATCTCAAGGAGAAAGTAATCTGTGTGATGTGATCTCAACC	429
Db	7	caggagaaaataaggaatgtcaagagaaaagtaattctgtcaatgtgtgtctcaact	66
QY	430	GTGGTTCGTCTTTTGGGAATATGTCAACAG-----	461
Db	67	gtgatcatgtgttttgggaatttatcaacagcacagaaggctcttcttctgtgatatat	126
QY	462	-----AATTCCAGAGTGTGTGAGACAGATGCGAAGGACTGGTGGTTCACAGC	513
Db	127	cactcaaaaaaacaggaagttgatgacagcagtcgccaagaaggctggtggtctctgagc	186
QY	514	TGGTTTAAATGGGACCCACAGTTTATCAAGAGACAAGCTAGAGGACGGAGAAAG	573

Db 187 tggtttaacaatggatcccaatattatcaacaagggaagagacacatagacaaagaaaa 246  
QY 574 GGTAGA-----AATGGAGATGCGCATTTGAAGAGCCTCAGCTATGGGACTGGT 619  
Db 247 ggaagagagagagacccaaaggaaggaatgacacacagagcttcggtatgggactggt 306

QY 620 TCAATCCAAA 629

Db 307 ttaatccaaa 316

## RESULT 2

US-10-115-123-119  
; Sequence 119, Application US/10115123  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30APID2  
; CURRENT APPLICATION NUMBER: US/10/115,123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1998-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 119  
; LENGTH: 1193  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-115-123-119

Query Match 5.7%; Score 85.4; DB 6; Length 1193;  
Best Local Similarity 64.1%; Pred. No. 4.2e-15;  
Matches 189; Conservative 0; Mismatches 56; Indels 50; Gaps 2;

QY 385 AATGTCAGGGAAGTAATCCCTGTTGATGCTGCTGCTCAACCGTGGTTCGCTGTTT 444

Db 68 aatgtaaggaagaaatctgtcaatgctgtctcaactgtgattgttt 127

QY 445 TGGGAATATGTCACAG-----AATTCCA 468

Db 128 tgggaattatcacagacagaaagagctcttcttgatatatactcaaaacccca 187

QY 469 GAGTGTGGTGAGAACAGATGGCAGAGGACTGGTGGTTCACAGCTGGTTTAAATAATGGG 528

Db 188 gaagttagacagacagctgctcagaaggctggtgttctgagctggtttaacaatggg 247

QY 529 ACCCAGATTATCAAGAAGACACACGTAGAGACGGAGAGAAAGGGTAGA----- 579

Db 248 atccacaattatacaaggggaagacacatagacaaagaaagagagagagacc 307

QY 580 -----AATGGAGATCCATTGAAGACCCCTCAGCTATGGGACTGTTCAATCCAAA 629

Db 308 aaaggaagaaaaatgacacacacagagcttcggtatgggactggtttaatcccaaa 362

## RESULT 3

US-09-540-210B-12481  
; Sequence 12481, Application US/09540210B  
; GENERAL INFORMATION:

; APPLICANT: Sellhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE  
; FILE REFERENCE: PD-1037 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,210B  
; CURRENT FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: 08/972,899  
; PRIOR FILING DATE: November 18, 1997  
; PRIOR APPLICATION NUMBER: 08/395,244  
; PRIOR FILING DATE: February 27, 1995  
; PRIOR APPLICATION NUMBER: 08/722,922  
; PRIOR FILING DATE: September 27, 1996  
; PRIOR APPLICATION NUMBER: 60/005,526  
; PRIOR FILING DATE: September 29, 1995  
; PRIOR APPLICATION NUMBER: 08/824,029  
; PRIOR FILING DATE: March 25, 1997  
; PRIOR APPLICATION NUMBER: 60/014,010  
; PRIOR FILING DATE: March 25, 1996  
; PRIOR APPLICATION NUMBER: 08/826,847  
; PRIOR FILING DATE: April 10, 1997  
; PRIOR APPLICATION NUMBER: 60/015,533  
; PRIOR FILING DATE: April 10, 1996  
; PRIOR APPLICATION NUMBER: 08/903,555  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/023,308  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/862,178  
; PRIOR FILING DATE: May 22, 1997  
; PRIOR APPLICATION NUMBER: 60/018,217  
; PRIOR FILING DATE: May 23, 1996  
; PRIOR APPLICATION NUMBER: 08/881,589  
; PRIOR FILING DATE: June 24, 1997  
; PRIOR APPLICATION NUMBER: 60/021,275  
; PRIOR FILING DATE: June 25, 1996  
; PRIOR APPLICATION NUMBER: 08/903,802  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/023,308  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/905,881  
; PRIOR FILING DATE: August 1, 1997  
; PRIOR APPLICATION NUMBER: 60/025,204  
; PRIOR FILING DATE: August 1, 1996  
; PRIOR APPLICATION NUMBER: 08/903,471  
; PRIOR FILING DATE: July 30, 1997  
; PRIOR APPLICATION NUMBER: 60/025,478  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/903,556  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/025,217  
; PRIOR FILING DATE: August 22, 1996  
; PRIOR APPLICATION NUMBER: 08/937,142  
; PRIOR FILING DATE: September 23, 1997  
; PRIOR APPLICATION NUMBER: 60/026,598  
; PRIOR FILING DATE: September 24, 1996  
; PRIOR APPLICATION NUMBER: 08/960,746  
; PRIOR FILING DATE: October 29, 1997  
; PRIOR APPLICATION NUMBER: 60/030,144  
; PRIOR FILING DATE: October 30, 1996  
; PRIOR APPLICATION NUMBER: 08/826,847  
; PRIOR FILING DATE: April 10, 1997  
; PRIOR APPLICATION NUMBER: 60/015,533  
; PRIOR FILING DATE: April 10, 1996  
; PRIOR APPLICATION NUMBER: 08/755,524  
; PRIOR FILING DATE: November 22, 1996  
; PRIOR APPLICATION NUMBER: 60/007,495  
; PRIOR FILING DATE: November 22, 1995  
; PRIOR APPLICATION NUMBER: 09/021,031  
; PRIOR FILING DATE: February 10, 1998  
; PRIOR APPLICATION NUMBER: 60/039,325

```
; PRIOR FILING DATE: February 13, 1997
; PRIOR APPLICATION NUMBER: 09/035,172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431
; PRIOR FILING DATE: May 29, 1997
; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35654
; SOFTWARE: PERL Program
; SEQ ID NO 12481
; LENGTH: 305
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00676781
US-09-540-2108-12481

Query Match          5.6%; Score 84.2; DB 5; Length 305;
Best Local Similarity 56.8%; Pred. No. 5.3e-15;
Matches 155; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 1119 CGAGAAGTTCACCTATGAGAGCGGGAAGTCTCGCGCGGTACATTCCTCCAGGAGGG 1178
Db 2 ccagcagctccccctgagcgcaggcgtgttccactgcttctgtgagacagcgaagg 61

QY 1179 GGATTTTACTACCGCGCCATTTTGGAGGAACGCCCTACTCACATTCTCAACCTCAC 1238
Db 62 ggaactctattatggttggggcagctcttcggggggcgaggtgcccaggggtatatgagtttac 121

QY 1239 CAGGGAGTGTCTTTAAGGGGATCTCCAGGACAAAGAACATGACATAGAACCCCGAGTGGCA 1298
Db 122 taggggctgccacatggccatctctggcgacaaaggccaatggcatcatggtcgtcgctggcg 181

QY 1299 TGATGAGAGCCACTCAACAATACHTCTCTTTCAACAACCCCACTAAAATCCTATCTCC 1358
Db 182 gggaggaaagccactgaaccgtcactctcatctcaaaagccgctccaaggtgctgtcccc 241

QY 1359 AGAGTATGCTGGGACTATCAGATAGGCCCTGCC 1391
Db 242 cgagtacctctgggagcaggaagcccccagcc 274

RESULT 4
US-10-102-524-1730
; Sequence 1730, Application US/10102524
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; THERAPY AND DIAGNOSIS OF KIDNEY CANCER
```

```
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1730
; LENGTH: 3186
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-1730

Query Match          2.5%; Score 37.8; DB 6; Length 3186;
Best Local Similarity 48.8%; Pred. No. 0.59;
Matches 102; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 941 AGGATATCAGCATGATGCCCATCAAGACCATTGGGGAGGACACATCCTGGCCCATCCAGC 1000
Db 193 aggacatgagcaaggagcaggttgaggagcatgtccgcgcacatccggggagagctggacc 252

QY 1001 ACGAGGTGCACTTCCTCTTCTGTCATGGAGCTCAAGTCTTTCAAGACAACCTTCGGGG 1060
Db 253 ggcagcggggaggagaaactactctccagctggagcgggacaaagatccacaccttctggg 312

QY 1061 TGGAAACTCTGGGCCAGCTGGTAGACACAGCTCCAGGCCCTGGTGTACAAAGGCCAGTCCCG 1120
Db 313 agatcacagcggagcagctggaggagaagagctgagctgcggaaacaaagaccgggaga 372

QY 1121 AGAAGTTCACCTATGAGAGCGGGAACCTG 1149
Db 373 tggagaagccggaggagagccaccaggtg 401

RESULT 5
US-09-053-375B-710/c
; Sequence 710, Application US/09053375B
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: Nucleic Acid Arrays
; FILE REFERENCE: CLON-006
; CURRENT APPLICATION NUMBER: US/09/053,375B
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 1543
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 710
; LENGTH: 8809
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-053-375B-710

Query Match          2.5%; Score 36.8; DB 5; Length 8809;
Best Local Similarity 80.0%; Pred. No. 1.8;
Matches 56; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

QY 309 CAAGATCTCCATGTCGAAGTCCAGATCCAGTCCAGAACAAAG--TCTTCCATCCTCAAGATCTGGA 366
Db 8788 CAAGATCTCCATACCAAGATCCAGGTCCAGGTCAGAAACTTGTATCTCCAGGCTCCAGATTAGGA 8729

QY 367 TCACAGGAGA 376
Db 8728 TCACTGGTGA 8719

RESULT 6
US-10-123-155-204/c
; Sequence 204, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
```

```

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: ACIDS ENCODING THE SAME
; CURRENT APPLICATION NUMBER: P3330R1C30
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 204
; LENGTH: 790
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-123-155-204

Query Match      2.4%  Score 36.2; DB 6; Length 790;
Best Local Similarity 8.5%; Pred. No. 0.98;
Matches 35; Conservative 130; Mismatches 248; Indels 0; Gaps 0;

Qy 749 CAGTGTGTTGCTGGGAAAGTACATTGAGCATTTCTAGAGACTTTCTGGAGCTCTG 808
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 483 MRBHN.GSSSNCHGR.T.HG.NC..YGBRCS.Y.YGHDGG.C.SGNTC...MGNS.GGYH 424

Qy 809 ACATGTACTCATGGTGGCCATCGGTCATATTTACGTGATCATGATAGACACACTCC 868
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 423 M.G..ASCGHG..NH.B.MNCGYM.BRC.C..NN.GATCSS.DM...S.RSGNS.CAYCSC 364

Qy 869 GGATCGCTGCTGCTGACCTGAACCTCTACATTCCTTACAACTCTTTGAGATCAGGTCG 928
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 363 K.GN.YCYG.GR..HDSKBTDC..DTSWAT.T..NAH...ASBSCTR..SGBMC.SWT.T 304

Qy 929 AGAAGAGTGGCAGGATATCAGCATGATGCGCATGAGACCATTTGGGAGACATCTGG 988
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 HNRMY..R.DTWSS.YTN.BSMNR.B.YC.B.T...RCBY.MN.HAWB..TM.D.YKYB 244

Qy 989 CCCACATCCAGCAGGTCGACTCTCTCTGATGCGAGCTGATCAAGTCTTTCAAG 1048
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 H..RAHK.CA...THD.B.S.NNMB.YRYS.H.....RM.DMRSC..HY.YTMN.R.T 184

Qy 1049 ACAACTTGGGGTGAACACTCTGGCCAGCTGGTGGACACAGCTCCAGGCTGGGTACA 1108
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 183 K.W...HHS.CGB.NC...MM..RBBD..S.STM.....RDMTH.N.BCY.CCKGAS.AT 124

Qy 1109 AGGCCATCCCGAGAAGTTCACCTATGAGAGCGGGGAACATGTCGGCGCGCTAC 1161
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 MSH.S.MBSCKRNGHM..R.DH....CD.A.S.BY.DY.....YM.W.D.BDMM 71

RESULT 7
US-09-647-041C-7
; Sequence 7, Application US/09647041C
; GENERAL INFORMATION:
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha
; APPLICANT: McConnell, Stephen J.
; APPLICANT: Spinella, Dominic G.
; TITLE OF INVENTION: PEPTIDE LIGANDS FOR THE HUMAN FIBROBLAST GROWTH FACTOR (FGF) RECEPTOR
; FILE REFERENCE: CH005-01.US
; CURRENT APPLICATION NUMBER: US/09/647, 041C
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/11844
; PRIOR FILING DATE: 1999-05-28

```

```

; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc_feature
; LOCATION: (1)..(837)
; OTHER INFORMATION: Synthetic
; US-09-647-041C-7

Query Match      2.4%  Score 36.2; DB 5; Length 837;
Best Local Similarity 49.8%; Pred. No. 1;
Matches 119; Conservative 0; Mismatches 118; Indels 2; Gaps 1;

Qy 669 GCCGATTGTGTGGGAGGCACTTATGACACAGCTCTGCTGGAAAAGTACTACGCCACACA 728
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 260 gcccttcagtgaggaccccaaacccacactgcgtggttgaaaaatggcaagaattca 319

Qy 729 GAAACTCAGTGTGGGCTGACAGTGTGTTGCTGTGGGAAAGTACATTTGAGCATTTACTT 788
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 320 aacctgaccacagaaattggagggtacaagggtccggttatgccacctggagcat--cataat 377

Qy 789 AGACTTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGTCATATTTTACGT 848
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 378 ggaactcgtgtggtgcccctctgacaagggtgacacactacacactgcatgtgagaaatgagtcagg 437

Qy 849 CATGATAGACACACACTCCCGGATGCCCTGCTGTCGACCTGAACCTCTACATTTCTTAC 907
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 438 cagcatcaaccacacataccagctggatgctgtgagcgggtccctccaccggcccatcc 496

RESULT 8
US-09-053-375B-494
; Sequence 494, Application US/09053375B
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: Nucleic Acid Arrays
; FILE REFERENCE: CLON-006
; CURRENT APPLICATION NUMBER: US/09/053, 375B
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 1543
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 494
; LENGTH: 3328
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-053-375B-494

Query Match      2.4%  Score 36.2; DB 5; Length 3328;
Best Local Similarity 49.8%; Pred. No. 1.8;
Matches 119; Conservative 0; Mismatches 118; Indels 2; Gaps 1;

Qy 669 GCCGATTGTGTGGGAGGCACTTATGACACAGCTCTGCTGGAAAAGTACTACGCCACACA 728
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 488 gcccttcagtgaggaccccaaacccacactgcgtggttgaaaaatggcaagaattca 547

Qy 729 GAAACTCAGTGTGGGCTGACAGTGTGTTGCTGTGGGAAAGTACATTTGAGCATTTACTT 788
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 548 aacctgaccacagaaattggagggtacaagggtccggttatgccacctggagcat--cataat 605

Qy 789 AGACTTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGTCATATTTTACGT 848
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 606 ggaactcgtgtggtgcccctctgacaagggtgacacactacacactgcatgtgagaaatgagtcagg 665

Qy 849 CATGATAGACACACACTCCCGGATGCCCTGCTGTCGACCTGAACCTCTACATTTCTTAC 907
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



Db 666 cagcatcaaccacataccagctggatgtctgtgagcgggtccctccatccacccggccatcc 724

RESULT 9  
US-10-123-155-358  
; Sequence 358, Application US/10123155  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330RIC30  
; CURRENT APPLICATION NUMBER: US/10/123,155  
; CURRENT FILING DATE: 2002-04-15  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-123-155-358

Query Match 2.4%; Score 36; DB 6; Length 1049;  
Best Local Similarity 6.6%; Pred. No. 1.3;  
Matches 31; Conservative 121; Mismatches 320; Indels 0; Gaps 0;  
QY 645 TTTCACAGTACCCCGTGAAGCGCCGATTTGTGGGAGGACCTTATGACACAGCTCT 704  
Db 23 KILGARWFKPLPCDVTLDVPRNHVIVDCTDKHLTIPGGIPTNTNLTLTINHIPDISP 82  
QY 705 GCTGGAAGAGTACTACGCCACACAGAACTCACTGTGGGGCTGACAGCTGTTCCTGTGG 764  
Db 83 ASFHRLDHLVEIDFRNCVPIPLGSKNNMCIKRLQIKRFSGLTYLKLSDLDGNOLLEI 142  
QY 765 AAAGTACATTGACCAFTACTTAGAAGACTTTCTGGAGTCTGCTGACATGACTTTCATGCT 824  
Db 143 PQLPSPQLLSLEANNIFSIKENITELANIELYLQNCYYRNPVYSYSIERDAFLN 202  
QY 825 TGCCATCGGGTCATATTTAGTCATGATAGACAGCACTCCCGGATCCCTGTCTGCA 884  
Db 203 LTKLVLSLKDNNVTAVPTLPSTLTLEYLYNNMTAKIQEDDFNNINQLQLDLSGNCPR 262  
QY 885 CCGTGAACCTCTACATTCCTTACAGTCTTTGAGATCAGTCTGTGAGAAGAGGTGGCAGCA 944  
Db 263 CYNAPPAPCAPKNSPLQIPVNAFDALTELKVLRLHSNLSLQHVPPFRFNKLNQLQELDLS 322  
QY 945 TATCAGCATGATCGGCATGAAGACCACTTGGGAGCAGCATCTCGCCACATCCAGCACGA 1004  
Db 323 ONFLAKEIGDARFLHPLSLIQLDLSFNELQVYRASNLSQAFSLKSLKILIRIRGYVP 382  
QY 1005 GGTCGACTTCTCTTCGATGACGTGGATCAAGTCTTTCAAGACAACTTCGGGTGGA 1064  
Db 383 KELKSNLSPHLNQLNQLVDLTNEIKIANLSMFKQFRRLKVIDLSVKNISPSGDSSEV 442  
QY 1065 AACTCTGGCCAGCTGGTAGCAGCTCCAGGCGCTGGTGGTACAGGCCAGT 1116  
Db 443 GFCSNARTSVESYEPQVLEQLHYFRDYKARSRCFRKNKEASFSMVNESCKY 494

RESULT 10  
US-10-102-806-97  
; Sequence 97, Application US/10102806  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA103PIC1  
; CURRENT APPLICATION NUMBER: US/10/102,806  
; CURRENT FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 09/925,298  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05881  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 846  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 97  
; LENGTH: 658  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (627)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (634)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (635)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-102-806-97

Query Match 2.4%; Score 35.6; DB 6; Length 658;  
Best Local Similarity 47.0%; Pred. No. 1.4;  
Matches 110; Conservative 0; Mismatches 124; Indels 0; Gaps 0;  
QY 842 TTTCACATGATAGACGACACCTCCCGGATCCCTGTCGTGACACCTGAACCCCTCTACATT 901  
Db 40 tctacgcccgcactacaacatcccgggtgatccatccctccgcggcgctggagcacc 99  
QY 902 CCTTACAAGTCTTTGAGATCAGTCTGAGAAGAGGTGGCAGGATATCAGCATGATCGCA 961  
Db 100 ctggcctggtttcaaccagctgcccagatgctgtaccccagttaccacaagtgacc 159  
QY 962 TGAAGACCATTTGGGAGCAGACATCTGCGCCACATCAGCAGCAGGTCGACTTCTCTTCT 1021  
Db 160 agatgatgctggagcagtcctcctctgcccagcccctatgaggggttaccgagcctcc 219  
QY 1022 GCATGACCTGGATCAAGTCTTTCAAGACAACTTCGGGTGGAAACTCTCGGCC 1075  
Db 220 ccaggcaccagctgctgtgtcttcaaggagactgcccaggccgtgttccaggacc 273

RESULT 11  
US-10-041-018-399  
; Sequence 399, Application US/10041018  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seichi P.T.  
; APPLICANT: Hart, Elizabeth A.  
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
; FILE REFERENCE: P02080US1/10025547  
; CURRENT APPLICATION NUMBER: US/10/041,018  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259880  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 399  
; LENGTH: 39827



; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: full length human CLASP-7 cDNA
; NAME/KEY: CDS
; LOCATION: (13)..(6156)
; OTHER INFORMATION: human CLASP-7
; US-09-736-968A-1

Query Match 2.3%; Score 35.2; DB 5; Length 6372;
Best Local Similarity 55.8%; Pred. No. 4.6;
Matches 67; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 981 CATCTGCCCCACATCCAGCAGGTCGACTTCCTTCTCCATGGAGCGTGATCAAGT 1040
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2196 caccctggcgactcctggagaggagcctccattcccgctcaaggacacagtgtct 2255

QY 1041 CTTTCAAGACAACTTCGGGGTGGAACTCTGGCCAGCTGGTAGCACAGCTCCAGGCCTG 1100
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2256 gaggcagggaacagtggagcaggagctgcgggcccagctcttcgacgactgcgcctggccag 2315

RESULT 15
US-10-123-155-160
; Sequence 160, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 160
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-123-155-160

Query Match 2.3%; Score 34.6; DB 6; Length 605;
Best Local Similarity 8.5%; Pred. No. 2.6;
Matches 47; Conservative 135; Mismatches 371; Indels 0; Gaps 0;

QY 427 ACCGTGGTGTGCTGTTTGGGAATATGTCAACAGAAATTCACAGAGTTCAGAGTTCGAGAGCAGA 486
Db 35 DEPGPEGLTSTSLDLLLPTGLEPLDSEPTMTGLGAGLGLGASGSGFPSENEESRILQP 94
QY 487 TGGCAGAAGGACTGGTGTCCCAAGCTGGTTTAAATAATGGGACCCACACAGTTATCAGAA 546
Db 95 PQYFWEEEEELNDSSLDLPTADYVFPDLTERAGSTEDTSAQELPNLPSLPKMLNLEP 154
QY 547 GACAACGTAGAGGACGGAGAGAGAAAGGTAGAAATGGAGATCGCATTCAGACAGCCTCAG 606
Db 155 PHMPPREEEEEEEEEEEKEEVEKEEEEEELLPVNGSQQEAAKQVDRSLTSSSOT 214
QY 607 CTATGGGACTGGTTCATCAATCCAAAGAACCCCGGATGTTTTCACAGTACACCCCGTGAAG 666
Db 215 PCATKSRHEDSGDAQSSGVESSMGPSLLPSVPTTVPDQDSTSOEAEATVLPAAAG 274
QY 667 GCGCCGATTGTGGGAAGGCACCTTATGACACAGCTCTCTCTGAAAAGTACTACGCCACA 726
Db 275 LCVFEFAPOEASEEATAGAAAGLSGQHEEVPALPSPQTTPAPSGAEHPDEDPLGSRTSASS 334
QY 727 CAGAAACTCACTGTGGGGCTGACAGTGTGCTGTGGGAAAGTACATTCAGCATTA 786
Db 335 PLAPGDMELTPSATLQGEDLNQQLLEGOAAEQSRIPWDSTQVICKDHSNLAKNYIL 394
QY 787 GAAGACTTTCATGAGTCTGCTGACATGATGTTTTCATGTTGGCCATCGGCTCATATTTTAC 846
Db 395 NMTENIDCEVFRHGRGPQLLALVEEVLPRHSGHHGAWHISLKPSEKQHLMLTLVGREQ 454
QY 847 GTCATGATAGACGACACCTCCCGGATGCTGCTGCTGACACCTGACACCTCTACATTCCTTA 906
Db 455 GVPTQDVLMLGDIRRSLEEIGNYSTSSQAHASQVRSYDGTFLFVLVYVIGAI 514
QY 907 CAAGTCTTTGAGATCAGGTCTGAGAGAGAGTGGCAGGATATCAGCATGATGCGCATGAAG 966
Db 515 IIALGLLYNCWRRRLPKLKHVSHGBELRPFVNGCHDNPTLDVASDSQSEMEKHPSLNGG 574
QY 967 ACCATTGGGGAGC 979
Db 575 GALNGPGSMALM 587

Search completed: May 17, 2002, 23:59:16
Job time: 19113 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:07:12 ; Search time 13.25 Seconds  
(without alignments)  
1090.260 Million cell updates/sec

Title: US-09-863-475A-4  
Perfect score: 2136  
Sequence: 1 MITMLQDLHVKNISRSKS.....IKSVKVAQTKYLNLRNV 394

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36564827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2136	100.0	394	1	GATR_MOUSE
2	1510	70.7	371	1	GATR_PIG
3	1476.5	69.1	368	1	GATR_BOVIN
4	690.5	32.3	354	1	BGAT_HUMAN
5	97	4.5	569	1	GCLL_DROME
6	96.5	4.5	301	1	OMPG_ECOLI
7	95.5	4.5	361	1	Y439_METJA
8	95	4.4	505	1	AGAR_ALTAT
9	91.5	4.3	455	1	SYS_PYRAB
10	91.5	4.3	455	1	SYS_PYRHO
11	91	4.3	323	1	SYK3_PASMU
12	90.5	4.2	496	1	RECO_BACSU
13	90	4.2	1478	1	YAWB_SCHPO
14	88.5	4.1	2896	1	HCYG_OCTDO
15	88	4.1	393	1	KTRL_YEAST
16	88	4.1	534	1	FM02_CAVPO
17	87.5	4.1	1941	1	UBR1_KLOLA
18	87	4.1	865	1	SYV_THEMA
19	87	4.1	1102	1	YM66_YEAST
20	86	4.0	255	1	RL4_PYRAB
21	86	4.0	644	1	YEN9_YEAST
22	86	4.0	1258	1	ATB1_HUMAN
23	86	4.0	1380	1	SP14_YEAST
24	85.5	4.0	887	1	TOP1_BACAN
25	85.5	4.0	1541	1	MRP2_RAT
26	85	4.0	1220	1	ATB1_PIG
27	84.5	4.0	724	1	SECL_YEAST
28	84.5	4.0	730	1	PHLC_PSEAE
29	84.5	4.0	1826	1	SUIS_HUMAN
30	84	3.9	396	1	DCAM_YEAST
31	84	3.9	1258	1	ATB1_RAT
32	83.5	3.9	386	1	CVCB_PEA
33	83.5	3.9	406	1	SBMA_ECOLI

ALIGNMENTS

RESULT	ID	GATR_MOUSE	STANDARD	PRT	394 AA
AC	P23336				
DT	01-NOV-1991	(Rel. 20, Created)			
DT	01-NOV-1991	(Rel. 20, Last sequence update)			
DT	30-MAY-2000	(Rel. 39, Last annotation update)			
DE	N-ACETYLLACTOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE (EC 2.4.1.151)				
DE	(GALACTOSYLTRANSFERASE) (UDP-GALACTOSE:BETA-D-GALACTOSYL-1,4-N-ACETYL-				
DE	D-GLUCOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE).				
GN	GGTA1 OR GGTA-1.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90046769; Pubmed=2510162;				
RA	Larsen R.D., Rajan V.P., Ruff M.M., Kukowska-Latallo J.,				
RA	Cummings R.D., Lowe J.B.;				
RT	"Isolation of a cDNA encoding a murine				
RT	UDPgalactose:beta-D-galactosyl-1,4-N-acetyl-D-glucosaminide				
RT	alpha-1,3-galactosyltransferase: expression cloning by gene				
RT	transfer.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:8227-8231(1989).				
RP	[2]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=92184813; Pubmed=1544928;				
RX	Joziasse D.H., Shaper N.L., Kim D., van den Eljnden D.H., Shaper J.H.;				
RA	"Murine alpha 1,3-galactosyltransferase. A single gene locus				
RT	specifies four isoforms of the enzyme by alternative splicing.";				
RL	J. Biol. Chem. 267:5534-5541(1992).				
CC	-1- FUNCTION: TRANSFER OF GALACTOSE FROM UDP-GALACTOSE TO AN				
CC	ACCEPTOR MOLECULE (R).				
CC	-1- CATALYTIC ACTIVITY: UDP-GALACTOSE + BETA-D-GALACTOSYL-1,4-				
CC	N-ACETYL-D-GLUCOSAMINYL-R = UDP + ALPHA-D-GALACTOSYL-1,3-				
CC	BETA-D-GALACTOSYL-1,4-N-ACETYL-D-GLUCOSAMINYL-R.				
CC	-1- COFACTOR: MANGANESE.				
CC	-1- PATHWAY: GLYCOSYLATION.				
CC	-1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND				
CC	FORM IN TRANS CISTERNAE OF GOLGI.				
CC	-1- SIMILARITY: STRONG, TO BLOOD GROUP AB TRANSFERASE.				
CC	-1- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN				
CC	GLYCOSYLTRANSFERASES.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; M26925; AAA37657.1; -;				
DR	EMBL; M85153; AAA37711.1; -;				
DR	PIR; A34417; A34417.				

Q9un88 homo sapien  
Q78507 guillardia  
P11997 drosophila  
P26446 gallus gall  
O29622 archaeoglob  
Q91fl1 mus musculu  
Q9z4p7 streptococc  
O53175 mycobacteri  
Q58611 methanococc  
P28603 azospirillu  
P13377 trypanosoma  
Q9xt19 drosophila

DR MGD: MCI:95704; Ggta1.  
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;  
KW Signal-anchor; Golgi stack; Alternative splicing; Manganese.  
FT DOMAIN 1 41 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 42 60 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).  
FT DOMAIN 61 394 LUMENAL, CATALYTIC (POTENTIAL).  
FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 1 35 MISSING (IN SHORTER ISOFORM).  
FT VARSPLIC 62 62 R -> SPDGSLWYHTK (IN SHORTER ISOFORM).  
SQ SEQUENCE 394 AA; 46475 MW; 7766831640D1BBF7 CRC64;

Query Match 100.0%; Score 2136; DB 1; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.2e-173;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQGEKIMNKGKIVLLMLIVSTVVVFWYEV 60  
Db 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQGEKIMNKGKIVLLMLIVSTVVVFWYEV 60

Qy 61 NRIPEVGNRWQKDWFFSWFKNGTHSQEDNVEGRREKGRNGDRIEPQLWDWENPKNR 120  
Db 61 NRIPEVGNRWQKDWFFSWFKNGTHSQEDNVEGRREKGRNGDRIEPQLWDWENPKNR 120

Qy 121 PDVLTVPWKAPIVWEGYDTALKEKYATOKLTGLTVFAVGKYIEHYLEDLESADMY 180  
Db 121 PDVLTVPWKAPIVWEGYDTALKEKYATOKLTGLTVFAVGKYIEHYLEDLESADMY 180

Qy 181 FMVGHVIFYVYIMDDTSRMPVHNLPLSLQVFEIRSEKRWQDISMMRMKTIGEHLAHI 240  
Db 181 FMVGHVIFYVYIMDDTSRMPVHNLPLSLQVFEIRSEKRWQDISMMRMKTIGEHLAHI 240

Qy 241 QHEVDFLCMDVDQVDFQNFVETLGLVAQLQAWWYKASPEKFTYERRELSAAVIFPGE 300  
Db 241 QHEVDFLCMDVDQVDFQNFVETLGLVAQLQAWWYKASPEKFTYERRELSAAVIFPGE 300

Qy 301 GDFYHAAIFGTPPHILNLTRFCKGLTQDKKHDEAQAHDHSHLNKYFLFNKPTKILS 360  
Db 301 GDFYHAAIFGTPPHILNLTRFCKGLTQDKKHDEAQAHDHSHLNKYFLFNKPTKILS 360

Qy 361 PEYCDWYQIGLPSDIKSVKVAWQKEYNLVRNV 394  
Db 361 PEYCDWYQIGLPSDIKSVKVAWQKEYNLVRNV 394

RESULT 2  
GATR\_PIG STANDARD; PRT; 371 AA.  
AC P50127;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE N-ACETYLLACTOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE (EC 2.4.1.151)  
DE (GALACTOSYLTRANSFERASE) (UDP-GALACTOSE:BETA-D-GALACTOSYL-1,4-N-ACETYL-  
DE D-GLUCOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE).  
GN GGTAL.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Sulina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YORKSHIRE; TISSUE=Spleen;  
RX MEDLINE=95104914; PubMed=7528726;  
RA Strahan K.M., Gu F., Preece A.F., Gustavsson I., Andersson L.,  
RA Gustafsson K.;  
RT "cDNA sequence and chromosome localization of pig alpha 1,3  
RL galactosyltransferase.";  
RN Immunogenetics 41:101-105(1995).  
RP [2]  
RC TISSUE=Spleen, and Liver;

RA Sandrin M.S., Dabkowski P.L., Henning M.M., Moutouris E.,  
RA McKenzie I.F.C.;  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
CC !- FUNCTION: TRANSFER OF GALACTOSE FROM UDP-GALACTOSE TO AN  
CC ACCEPTOR MOLECULE (R).  
CC !- CATALYTIC ACTIVITY: UDP-GALACTOSE + BETA-D-GALACTOSYL-1,4-  
CC N-ACETYL-D-GLUCOSAMINYL-R = UDP + ALPHA-D-GALACTOSYL-1,3-  
CC BETA-D-GALACTOSYL-1,4-N-ACETYL-D-GLUCOSAMINYL-R.  
CC !- COFACTOR: MANGANESE.  
CC !- PATHWAY: GLYCOSYLATION.  
CC !- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND  
CC FORM IN TRANS CISTERNAE OF GOLGI.  
CC !- SIMILARITY: STRONG, TO BLOOD GROUP AB TRANSFERASE.  
CC !- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN  
CC GLYCOSYLTRANSFERASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L36152; AAA73558.1; -;  
KW EMBL; L36535; AAA58775.1; -;  
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;  
KW Signal-anchor; Golgi stack; Alternative splicing; Manganese.  
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 7 22 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).  
FT DOMAIN 23 371 LUMENAL, CATALYTIC (POTENTIAL).  
FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 27 39 SPEGSLEWYQSK -> R (IN SHORT ISOFORM).  
FT CONFLICT 227 227 M -> I (IN REF. 2).  
SQ SEQUENCE 371 AA; 43764 MW; CFC715B8D89993D4 CRC64;

Query Match 70.7%; Score 1510; DB 1; Length 371;  
Best Local Similarity 72.7%; Pred. No. 1.2e-120;  
Matches 271; Conservative 48; Mismatches 38; Indels 16; Gaps 4;

Qy 36 MNVKGKIVLLMLIVSTVVVFWYEVN-----RIPEVGNRWQKDWFFSWFKN 83  
Db 1 MNVKGKIVLLMLIVSTVVVFWYEVNINSPEGSLEWYQSKNPEVGSSA-QRGWFFPSWFN 59

Qy 84 GTHSY--QEDNVEGRREKGRNGDRIEPQLWDWENPKNRPDVLTVPWKAPIVWEGYDT 141  
Db 60 GTHSYHEDEDAIGNEKEQKEDNRGELP-LVDWFPNPKRPEVVTITRWKAPVWEGYTNR 118

Qy 142 ALLEKYATQKLTGLTVFAVGKYIEHYLEDLESADMYFMVGHVIFYVYIMDDTSRMPV 201  
Db 119 AVLNDNYAKQKLTGLTVFAVGKYIEHYLEEFISANTYFMVGHKVIFFIMVDDISRML 178

Qy 202 VHLNPLSLQVFEIRSEKRWQDISMMRMKTIGEHLAHIQHEVDVDFLCMDVDQVDFQNF 261  
Db 179 IELGPLRSKFVEIKSEKRWQDISMMRMKTIGEHLAHIQHEVDVDFLCMDVDQVDFQNF 238

Qy 262 VETLGLVAQLQAWWYKASPEKFTYERRELSAAVIFGEGDFYHAAIFGCTPTHTLNL 321  
Db 239 VETLGLVAQLQAWWYKASPEKFTYERRELSAAVIFGEGDFYHAAIFGCTPTHTLNL 298

Qy 322 RECFCGILQDKKHDEAQAHDHSHLNKYFLFNKPTKILSPEYCDWYQIGLPSDIKSVKVA 381  
Db 299 QECFCGILQDKKHDEAQAHDHSHLNKYFLFNKPTKILSPEYCDWYQIGLPSDIKSVKVA 358

Qy 382 WQTEYNLVRNV 394  
Db 359 WQTEYNLVRNV 371

RESULT 3  
GATR\_BOVIN

```
ID GATR_BOVIN STANDARD; PRT; 368 AA.
AC P14769;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE N-ACETYLGLUCOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE (EC 2.4.1.151)
DE (GALACTOSYLTRANSFERASE) (UDP-GALACTOSE:BETA-D-GALACTOSYL-1,4-N-ACETYL-
DE D-GLUCOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE).
GN GGTAL.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89340543; PubMed=2503516;
RA Jozlase D.H., Shaper J.H., van den Eljnden D.H., van Tunen A.J.,
RA Shaper N.L.;
RT "Bovine alpha 1->3-galactosyltransferase: isolation and
RT characterization of a cDNA clone. Identification of homologous
RT sequences in human genomic DNA."
RL J. Biol. Chem. 264:14290-14297(1989).
CC -|- FUNCTION: TRANSFER OF GALACTOSE FROM UDP-GALACTOSE TO AN
CC ACCEPTOR MOLECULE (R).
CC -|- CATALYTIC ACTIVITY: UDP-GALACTOSE + BETA-D-GALACTOSYL-1,4-
CC N-ACETYL-D-GLUCOSAMINYL-R = UDP + ALPHA-D-GALACTOSYL-1,3-
CC BETA-D-GALACTOSYL-1,4-N-ACETYL-D-GLUCOSAMINYL-R.
CC -|- COFACTOR: MANGANESE.
CC -|- PATHWAY: GLYCOSYLATION.
CC -|- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -|- DISEASE: AUTOIMMUNE DISEASE (ANTIBODIES AGAINST ENZYMATIC
CC PRODUCT).
CC -|- SIMILARITY: STRONG, TO BLOOD GROUP AB TRANSFERASE.
CC -|- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN
CC GLYCOSYLTRANSFERASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J04989; AAA30558.1; -
DR PIR; A44785; A44785
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack; Manganese.
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 7 22 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 23 368 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 368 AA; 43246 MW; 5BC50D6737BDDC33 CRC64;
Query Match 69.1%; Score 1476.5; DB 1; Length 368;
Best Local Similarity 72.2%; Pred. No. 8.5e-118;
Matches 268; Conservative 41; Mismatches 47; Indels 15; Gaps 3;
QY 36 MNYKGVILLMLIVSTVVVFWFVYV-----NRIPYGENRWQDWFPSWFKN 83
:|||||:|||||:|||||:
DB 1 MNYKGVILMSLVSTVIVVFWFYIHSPEGLFWINPSRNPVEVGGSSIQKGLWLPWFNN 60
QY 84 GTIYSQEDNVGREGKRGNDRIEELPOLDFNPKRNPDLVTPKAPVWEGTVDLAL 143
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 61 GYR--BEDGDINEEKEQRNDE-SKLKLSDFNPFKRPPEVVTWKAPVWEGTYNRAV 117
QY 144 LKRYATOKLTGLTVFVAVGKYIEHYLEDPLESADMYEMVGHVIFVIMDDTSRMPVVH 203
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 118 LDNYAKQKITVGLTVFVAGRYIEHYLEELFSLTSANKHFVGHVPIFYIVDDVSRMPLIE 177
```

```
QY 204 LNPHLSQLVEFIRSEKRWQDISMRMKTIGEHILAHQIHEVDFLFCMDVDVQVQDNFGVE 263
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 178 LGLRSLFKVFKIKPEKRWQDISMRMKTIGEHIVAHQIHEVDFLFCMDVDVQVQDNFGVE 237
QY 264 TLQQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGEGDFYHAAIFGGTPTTHLNLTR 323
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 238 TLGESVAQLQAWMYKADPNDFTYERRKESAAAYIPFGEGDFYHAAIFGGTPTQVNLNTOE 297
QY 324 CFKGILODKKHIDEAQWHDHSHLNKPTKILSPYCWQDYQICLPSDIKSVKVAWQ 383
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 298 CFKGILODKKHIDEAQWHDHSHLNKPTKILSPYCWQDYHIGLPADIKLVKMSWQ 357
QY 384 TKEYNLVRNV 394
DB 358 TKEYNLVRNV 368
RESULT 4
BGAT_HUMAN STANDARD; PRT; 354 AA.
AC P16442;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HISTO-BLOOD GROUP ABO SYSTEM TRANSFERASE (NAGAT) [INCLUDES:
DE GLYCOPROTEIN-FUCOSYLALACTOSIDE ALPHA-N-
DE ACETYLALACTOSAMINYLTRANSFERASE (EC 2.4.1.40) (FUCOSYLGLYCOPROTEIN
DE ALPHA-N-ACETYLALACTOSAMINYLTRANSFERASE) (HISTO-BLOOD GROUP A
DE TRANSFERASE) (A TRANSFERASE); GLYCOPROTEIN-FUCOSYLALACTOSIDE ALPHA-
DE GALACTOSYLTRANSFERASE (EC 2.4.1.37) (FUCOSYLGLYCOPROTEIN 3-ALPHA-
DE GALACTOSYLTRANSFERASE) (HISTO-BLOOD GROUP B TRANSFERASE) (B
DE TRANSFERASE)].
GN ABO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90110098; PubMed=2104828;
RA Yamamoto F.-I., Marken J., Tsuji T., White T., Clausen H.,
RA Hakomori S.-I.;
RT "Cloning and characterization of DNA complementary to human UDP-
RT GalNAc: Fuc alpha 1->2Gal alpha 1->3GalNAc transferase (histo-blood
RT group A transferase) mRNA."
RL J. Biol. Chem. 265:1146-1151(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90238543; PubMed=2333095;
RA Yamamoto F.-I., Clausen H., White T., Marken J., Hakomori S.-I.;
RT "Molecular genetic basis of the histo-blood group ABO system."
RL Nature 345:229-233(1990).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=91035461; PubMed=2121736;
RA Yamamoto F.-I., Hakomori S.-I.;
RT "Sugar-nucleotide donor specificity of histo-blood group A and B
RT transferases is based on amino acid substitutions."
RL J. Biol. Chem. 265:19257-19262(1990).
CC -|- FUNCTION: THIS PROTEIN IS THE BASIS OF THE ABO BLOOD GROUP SYSTEM.
CC THE HISTO-BLOOD GROUP ABO INVOLVES THREE CARBOHYDRATE ANTIGENS: A,
CC B, AND H. A, B, AND AB INDIVIDUALS EXPRESS A GLYCOSYLTRANSFERASE
CC ACTIVITY THAT CONVERTS THE H ANTIGEN TO THE A ANTIGEN (BY ADDITION
CC OF UDP-GALNAc) OR TO THE B ANTIGEN (BY ADDITION OF UDP-GAL),
CC WHEREAS O INDIVIDUALS LACK SUCH ACTIVITY.
CC -|- CATALYTIC ACTIVITY: UDP-N-ACETYL-D-GALACTOSAMINE + GLYCOPROTEIN
CC ALPHA-L-FUCOSYL-(1,2)-D-GALACTOSE = UDP + N-ACETYL-ALPHA-D-
CC GALACTOSAMINYL-(1,3)-[ALPHA-L-FUCOSYL-(1,2)]-D-GALACTOSE.
CC -|- CATALYTIC ACTIVITY: UDP-GALACTOSE + GLYCOPROTEIN ALPHA-L-FUCOSYL-
CC (1,2)-D-GALACTOSE = UDP + GLYCOPROTEIN ALPHA-D-GALACTOSYL-(1,3)-
CC [ALPHA-L-FUCOSYL-(1,2)]-D-GALACTOSE.
CC -|- PATHWAY: GLYCOSYLATION.
CC -|- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
```

CC FORM IN TRANS CISTERNAE OF GOLGI. SOLUBLE FORM IN BODY FLUIDS.  
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY  
CC PROTEOLYTIC PROCESSING.  
CC -1- POLYMORPHISM: THE SEQUENCE SHOWN IS THAT OF THE A TRANSFERASE. THE  
CC B FORM DIFFERS BY A FEW RESIDUES SUBSTITUTIONS, THE O PHENOTYPE IS  
CC RESULT OF A SINGLE BASE FRAMESHIFT DELETION IN THE N-TERMINAL  
CC EXTREMITY OF THE GENE.  
CC -1- SIMILARITY: STRONG. TO N-ACETYLGLUCOSAMINIDE ALPHA-1,3-  
CC GALACTOSYLTRANSFERASE (EC 2.4.1.151).  
CC -1- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN  
CC GLYCOSYLTRANSFERASES.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J05175; AAA36792.1; .  
CC DR PIR; A34933; A34933.  
CC DR PIR; S09593; S09593.  
CC DR MIM; I10300; .  
CC KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;  
CC KW Signal-anchor; Golgi stack; Polymorphism; Blood group antigen.  
CC FT FUGOSYLGLYCOPROTEIN ALPHA-N-  
CC CHAIN 54 354  
CC FT ACETYLGLACTOSAMINYLTRANSFERASE, SOLUBLE  
CC FT FORM.  
CC FT CYTOPLASMIC (POTENTIAL).  
CC FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
CC FT (POTENTIAL).  
CC FT LUMENAL, CATALYTIC (POTENTIAL).  
CC FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT CR -> W.  
CC FT /FTid=VAR\_003408.  
CC FT P -> L (IN ALLELE A2).  
CC FT /FTid=VAR\_003409.  
CC FT R -> G (IN GROUP B TRANSFERASE).  
CC FT /FTid=VAR\_003410.  
CC FT G -> S (IN GROUP B TRANSFERASE).  
CC FT /FTid=VAR\_003411.  
CC FT L -> M (IN GROUP B TRANSFERASE;  
CC FT IMPORTANT FOR THE SPECIFICITY).  
CC FT /FTid=VAR\_003412.  
CC FT G -> A (IN GROUP B TRANSFERASE;  
CC FT IMPORTANT FOR THE SPECIFICITY).  
CC FT /FTid=VAR\_003413.  
CC FT R -> W (IN ALLELE B3 OF GROUP B  
CC FT TRANSFERASE).  
CC FT /FTid=VAR\_003414.  
CC FT  
CC SQ SEQUENCE 354 AA; 40934 MW; A03DA16B630C1608 CRC64;  
Query Match 32.38; Score 690.5; DB 1; Length 354;  
Best Local Similarity 45.88; Pred. No. 3.3e-51;  
Matches 142; Conservative 47; Mismatches 108; Indels 13; Gaps 4;  
QY 95 GRREKG-----RNGDRIEPEQLWDFNPKN-----RPDVLTPVPKAPIVWEQTYDTAL 143  
DB 45 GSLRGFCMAVRPFDHQLQVSLPRVVPQKVLTPCKRDVLVTPWLPAPIVWEGTFNIDI 104  
QY 144 LEKYATQKLTGLTVFPAVGKYEIEHLEDFLESADMYFMVGHRRVIFYVMIDDTSRMPVH 203  
DB 105 LNEQFRLONTTIGLTVFAIKKYA-FLKLFLETAEKHFVGHVHYVFTDPAAPRVPT 163  
QY 204 LNPILSLQVEIRESEKRWQDSMRMTIGTGHILAHQHEVDFLCMDVDVDFODNFGVE 263  
DB 164 LGTRQLSVLEVRAYKRWQDSMRMERMISDFCERRFLSEVDYLVCVDVDFEFDHVGVE 223  
QY 264 TLGQLVAQLQAWMYKASPEKFTYRERLSAAYIPFGEGDFYHYHAAIFGGTPTTHLNLTR 323  
DB 224 ILTFLGTLHPGFGSSREAFYERRPQSQAYIPKDEGDFYLLGGFGGVSQVQVRLTRA 283

QY 324 CFKGILODKKHDIQAQWHDHSHLNKYLFLNKPTRKILSPYCWQYQ-IGLPSDIKSVKVM 382  
DB 284 CHQAMVQDQANGIEAVWHDHSHLNKYLRLHKPTKVLSPYELWQDQLLGWPAVRLKRLPTA 343  
QY 383 QTKENLVLRN 392  
DB 344 VPKNHQAVRN 353  
RESULT 5  
GCLL\_DROME STANDARD; PRT; 569 AA.  
AC Q01820: Q9VAX5;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE GRM CELL-LESS PROTEIN.  
GN GCL OR CG8411.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RX SEQUENCE FROM N.A.  
RP MEDLINE=92370679; PubMed=1380406;  
RA Jongens T.A., Hay B., Jan L.Y., Jan Y.N.;  
RT "The germ cell-less gene product: a posteriorly localized component  
RT necessary for germ cell development in Drosophila.";  
RL Cell 70:569-584 (1992).  
RN [2]  
RX SEQUENCE FROM N.A.  
RP STRAIN=BERKELEY;  
RC MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Arif J.F., Aghayani A., An H.-J., Andrews-pfannkuch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Butcham M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Klamell B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclik J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shier B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svrlkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195 (2000).



```
CC -I- FUNCTION: REQUIRED FOR THE SPECIFICATION OF POLE CELLS AND GERM
CC CELL FORMATION. MOTHERS WITH REDUCED GLC FUNCTION GIVE RISE TO
CC STERILE ADULT PROGENY THAT LACK GERM CELLS.
CC -I- SUBCELLULAR LOCATION: POLE PLASM.
CC -I- DEVELOPMENTAL STAGE: OOCYGENESIS AND EMBRYOGENESIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M97933; AAA28566.1; -.
DR EMBL: AE003836; AAF59048.1; -.
DR PIR: A43317; A43317.
DR FlyBase: FBgn0005695; gcl.
DR InterPro: IPR000210; BTB_POZ.
DR Pfam: PF00651; BTB; 1.
DR SMART: SM00225; BTB; 1.
DR PROSITE: PS0097; BTB; 1.
KW Developmental protein.
SQ SEQUENCE 569 AA; 65152 MW; A91A670C4B5F83A4 CRC64;

Query Match 4.5%; Score 97; DB 1; Length 569;
Best Local Similarity 21.3%; Pred. No. 1-2;
Matches 50; Conservative 42; Mismatches 81; Indels 62; Gaps 10;

QY 96 RREKGRNGDRI---EEPLQWDFNPKNRDPVLTVP--WKAPIWEGTYDPTAL----- 143
DB 19 RRRKRSTSSGLKDDPAQLDTPKPK-KLLTTTOYIKALFKKEKNSDVAVMALDKVM 77
QY 144 -LEKYVATQ-----KLTGLTVFAVGKVIHYLEDPLE 175
DB 78 HLKVLVSQSPYPTFNMGWRAEQNFIOITLDRITVASLDVFGSMYSDEIE--IE 135
QY 176 SADMYFVAGHRVIFYV--MTDDTSRPVHNLPLHSLQVFEI-----RSEKRWQD 223
DB 136 SADVISLATATFLHLDGIIDKCAEYVNDISPETAQYVEAAQYGVGVGKSTQWFQ 195
QY 224 ISMRMKTIGEHLAHIQHEVDFLCMDVD-QVFQDNFGVETLGOLVAQLQAWMY 277
DB 196 INLLSIYKQPNLLRHISIELMSALTASPDLYVMQTEFSLYTL-----LRTNMF 244

RESULT 6
ID OMPG_ECOLI STANDARD; PRT; 301 AA.
AC P76045;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE OUTER MEMBRANE PROTEIN G PRECURSOR.
GN OMPG OR B1319.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-35.
RC STRAIN=K12;
RX MEDLINE=98389658; PubMed=9721282;
RA Fajardo D.A., Cheung J., Ito C., Sugawara E., Nikaido H.,
RA Mlsra R.;
RT "Biochemistry and regulation of a novel Escherichia coli k-12 porin
RT protein, OmpG, which produces unusually large channels.";
RL J. Bacteriol. 180:4452-4459(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
```

```
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -I- FUNCTION: FORMS CHANNELS FUNCTIONALLY LARGER THAN THOSE OF
CC CLASSICAL PORINS.
CC -I- SUBUNIT: MONOMER.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U49400; AAC34720.1; -.
DR EMBL: AE000230; AAC74401.1; -.
DR EcoGene: EGI13428; OmpG.
KW Outer membrane; Transmembrane; Porin; Signal; Complete proteome.
FT SIGNAL 1 21
FT CHAIN 22 301 OUTER MEMBRANE PROTEIN G.
FT TRANSMEM 27 35 POTENTIAL.
FT TRANSMEM 47 57 POTENTIAL.
FT TRANSMEM 64 72 POTENTIAL.
FT TRANSMEM 89 98 POTENTIAL.
FT TRANSMEM 104 112 POTENTIAL.
FT TRANSMEM 129 136 POTENTIAL.
FT TRANSMEM 149 158 POTENTIAL.
FT TRANSMEM 172 182 POTENTIAL.
FT TRANSMEM 186 195 POTENTIAL.
FT TRANSMEM 201 209 POTENTIAL.
FT TRANSMEM 213 222 POTENTIAL.
FT TRANSMEM 230 238 POTENTIAL.
FT TRANSMEM 240 248 POTENTIAL.
FT TRANSMEM 254 265 POTENTIAL.
FT TRANSMEM 269 279 POTENTIAL.
FT TRANSMEM 289 300 POTENTIAL.
SQ SEQUENCE 301 AA; 34913 MW; 63D1234F3DCCFECF CRC64;

Query Match 4.5%; Score 96.5; DB 1; Length 301;
Best Local Similarity 22.4%; Pred. No. 0-57;
Matches 73; Conservative 52; Mismatches 118; Indels 83; Gaps 21;

QY 72 OKDMWPPSWFKNGTHSYQEDNVGRRKGRNGORIEEPQLWDFNPKNRDPVLTVPWKA 131
DB 24 RNDWHF-----NIGAMYEIENVEG---YGEDMDGLAEPVS--VFNAAN-----GPWRI 66
QY 132 PIWV--EGTYDTALLEK--YYATOKLTVGLTVFAVGKVIHYLEDPLESADMYFMV--GH 185
DB 67 ALAYIOGSPVDYSAGKRGTWDRPELEV-----HY---QPLENDDFSGLTGGF 112
QY 186 RVIFYVMID-----DTSRPVHNLPLHSLQVFEIRSEKRWQDISMMR-----MKTIGEHIL 237
DB 113 RNYGYHYVDEPGDKTANKQWKIAPDWDVKLTDDLRENGW--LSMYKFAANDLNTTG-YAD 169
QY 238 AHIOHEVDFLFCMDVDQVFDNFGVETLGOLVAQLQAWMYKASPEKTYERRELSAAYIP 297
DB 170 TRVETETGLQYFNETVALRVNYLE-----RGFNMDDSRNNGEFSTQEI-RAYLP 219
QY 298 FGSGDFYHYHAAIFGGTPTHTLNLTRECFCGLQDKKHIDIAQWIDESHNLKYLFNKPTK 357
DB 220 LTIGN---HSV---TPYTRIGLDR-----WSNMDWQDDIEREGHD---FNRVGLF----- 260
QY 358 ILSPEYCWYQIGLPSDYSKVKVWQ 383
DB 261 -----YQYDFQNGLSVSLE-YAFEWQ 280
```

```

Query Match      4.4%  Score 95;  DB 1;  Length 505;
Best Local Similarity 19.3%;  Pred. No. 1.5;
Matches 90;  Conservative 66;  Mismatches 137;  Indels 174;  Gaps 26;

Qy  41  KVLILLIIVTVVVVFVEYVNRIPV-----GENRWQKDWFPSPFKN--GTHS 87
      |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db  3  KVIPLLLVTSLSVAI-PTYIHTATTEVVVNLNVKHSVEGKSEFER-----KNHIKLS 53

Qy  88  YOEDN-VEGRREK-----GR-NG-----DRIEEPOLWDFNPKN---RPDV 123
      ||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db  54  TLNDNDWQGEEDKLIKMYMEELDVYFGRNDGGTVWNFNQAIEDPANICYADPNIARGQA 113

```

```

Query Match      4.3%  Score 91.5;  DB 1;  Length 455;
Best local Similarity 20.1%;  Pred. No. 2.6;
Matches 87;  Conservative 62;  Mismatches 139;  Indels 145;  Gaps 25;

Qy 10 VNKISMRSKRSKSETSLPSSRSQSQKIMNVGKVIKLLMLIVTVVFWFVNRIPVEGEN 69
    ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 47 INKLHERHNKIAIEI-----GKRRK---KGEVPELLAKSKEIVKRIESLEK--EVEEL 95

Qy 70 RWQKDWFPSPW-FKNGTH-----SYQEDNV-----EGRRK-----GRNDR 105
    : : : : : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 96 KKKIDYLYL-WRLPNITHPSVIGSEEDNPVIRFGKARWVEGHLERFLEQSOGK---- 149

Qy 106 IEEPQLMDWFNPKNRPDLVLTTPWKAPIVMBGTVDYDTALLEK-----YYATQKLTIV--G 156
    ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 150 -MEYEVLEW-RPKLHVLDLLETL-----GGADFARAARKVSGSRFYVLLNEIVILDLA 198

Qy 157 LTVFAVKYIE-----HYLEDFESADMYFMVGHRTVIFYWIDTSMRPPVHLNPL 207
    ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 199 LIRFALDELIKKGTTVPVPPVMMRRFVEEGSTTFEDEDVLYKVEGEDVLYLPTAE-HPL 257

Qy 208 HSLQVFEIRSEKQDQISMMRMKKTIGEHILAHIQHEYDFLFCMDVDQVQDFQDNFVETLQG 267
    ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 258 AGMHANETLDGK----DLPLLVY-----GTSPCRKKAGTAGKTKGIFRVHO 301
    ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

```

```

Db 268 LVALDQAWWYKASPEKFTYERKRELSAAYIPGEGDFYTHAAIFGTT-----PHIL 318
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
302 -----FHKV--EQFVSRPEES-----WEWHERIIIRNABELFOKLEIPRVV 341
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 319 NLTRCF--KCILQDKKHDTAEQWHDHSHLNKLYFLFNKPTKILSPVCWDYQVIGLPSDIK 376
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 342 NI-----CTGDJGVAAKYDIEAWMPGGRFR-----EVVSASNCTD----- 379
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 377 SVKVAWOTKEYNL 389
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 380 -----WQARRLNI 387
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
SYS_PYRHO
ID SYS_PYRHO STANDARD; PRT; 455 AA.
AC OS8441;
DT 15-DEC-1998 (Rel. 37, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SERYL-TRNA SYNTHETASE (EC 6.1.1.11) (SERINE--TRNA LIGASE) (SERRS).
GN SERS OR PH0710 OR PHCF014.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -1- CATALYTIC ACTIVITY: ATP + L-SERINE + TRNA(SER) =
CC AMP + DIPHOSPHATE + L-SERYL-TRNA(SER).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC

```

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; AP000003; BAA29801.1; AUT\_INIT.  
DR InterPro; IPR002106; AA\_trna\_ligase\_II.  
DR InterPro; IPR003364; Seryl\_trna\_N.  
DR InterPro; IPR002314; trna-synt\_2b.  
DR InterPro; IPR002317; trna-synt\_ser.  
DR Pfam; PF02403; Seryl\_trna\_N; 2.  
DR Pfam; PF00587; trna-synt\_2b; 2.  
DR PRINTS; PR00981; TRNASYNTHSER.  
DR PROSITE; PS00179; AA\_TRNA\_LIGASE\_II\_1; 1.  
DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; FALSE\_NEG.  
KW Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Complete proteome.  
SQ SEQUENCE 455 AA; 53255 MW; B8D8A418A0651710 CRC64;

Query Match 4.3%; Score 91.5; DB 1; Length 455;  
Best Local Similarity 21.1%; Pred. No. 2.6;  
Matches 91; Conservative 57; Mismatches 141; Indels 143; Gaps 26;

QY 10 VNKISMSKSKSETSLPSSRSGSOEKIMNVKGVILLMLIVSTVVVWFVEYVNR-----PE 65  
DB 47 INLRHERNKAIVEI-----GKRRK-----KGEVDLLAKSR-----EIVKRIGELENE 91  
QY 66 VGENRWQKDMFPSPW-FKNGTH-----SYQEDNV-----EGRREKGRNGDRIEPEQ-- 110  
DB 92 VEELKKKIDYVL--WRLPNITHPSVPVCKDENDNVPIRFWCKARVKWKGHLERFLEQSOGK 149  
QY 111 -----LWDFNFKNRDPDVLVTTPWKAIPVWEGTYDTALLEK-----YYATOKLTV---GL 157  
DB 150 MEYELEW-KPKLHVDLEIL-----GGADPARAAKVGSGRFYLLNEIVILDIAL 199  
QY 158 TVEAVGKVE-----HYLEDLESADMYVMVGHVIEYVMVIMDDTSRMPVHNLPLH 208  
DB 200 INFALDRLEIKGFTPIPPYVVRREVEGSGTSFEDFVIYKVEDLILYLIPTAB-HPLA 258  
QY 209 SLOWFEIRSEKRWQDISMRMKTIGEHILAHQHEVDFLCMDVDQVDFQDFGVTLOOL 268  
DB 259 GMAHEILDGK---DLPLLY-----GVSPCRKEAGTAGCKTGKIFRVHQ- 301  
QY 269 VAQLQAWYKASPEKFTYERRELSAAYIPFGEGDFYHYAAIFGGT-----PHTLN 319  
DB 302 -----FHKV--EQFVYSRPEES-----WEWHEKIRNAEELFQLEIPYRVVN 342  
QY 320 LTRECF--KGILQDKKHDEAQWHDHSHLNKYLFLNFKPTKILSPCYCWDYQIGLPSDIKS 377  
DB 343 I---CTGDLGYVAAKKYDIEA-WMPGQ-----GKFEVVVSANCTD----- 379  
QY 378 VKVAMQTKENYL 389  
DB 380 ----WQARRLNI 387

RESULT 11  
SYK3\_PASMU STANDARD; PRT; 323 AA.  
AC P57824;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PUTATIVE LYSYL-TRNA SYNTHETASE (EC 6.1.1.6) (LYSTINE--TRNA LIGASE)  
DE (LYSRS) (GX).  
DE POXA OR PM0202.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PM70;  
RX MEDLINE=21145866; PubMed=11248100;

RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida pm70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
CC -!- FUNCTION: COULD BE A LYSYL-TRNA SYNTHETASE.  
CC -!- CATALYTIC ACTIVITY: ATP + L-LYSINE + TRNA(LYS) = AMP +  
CC PYROPHOSPHATE + L-LYSYL-TRNA(LYS).  
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; AE006054; AAK02286.1; -  
DR InterPro; IPR002106; AA\_trna\_ligase\_II.  
DR InterPro; IPR002309; trna-synt\_2.  
DR Pfam; PF00152; trna-synt\_2; 1.  
DR PROSITE; PS00179; AA\_TRNA\_LIGASE\_II\_1; 1.  
DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; 1.  
KW Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Complete proteome.  
SQ SEQUENCE 323 AA; 36859 MW; 84AFE23F83417888 CRC64;

Query Match 4.3%; Score 91; DB 1; Length 323;  
Best Local Similarity 20.6%; Pred. No. 1.8;  
Matches 42; Conservative 30; Mismatches 80; Indels 52; Gaps 9;

QY 212 VFEIRSEKRWQDISMRMKTIGEHILAHQH-----EVDFLCMDVDQVDFQDNFG 261  
DB 1 MFE---QENWQSPASIEENLLARAKIIAEIRRFETGRLLEVETPVLSFGVTDVHLSTFN 57  
QY 262 VETLGOLVAQLQAWYKASPEKFTYERRELSAAYIPFGEGDFYHYAAIFGGTPTTHLNT 321  
DB 58 TTFISPTAEKSKALWLSPE---YHMKRLAA-----GSGPIFQLCHVFRNEA----- 104  
QY 322 RECFCGILQDKKHDE---AOHWDHSHLNKYLFLN-----KPTKILSPCYCWD 366  
DB 105 -----GQHNPEFTMLEWY-RPHFDMYRLINEVDDLQOQILQCKPTESLSQFVQ 154

QY 367 YQIGL-PSDIKSVKVAWQTKENYL 389  
DB 155 EVVGLDPLSAEKAELVAKAKQVHL 178  
RESULT 12.  
RECO\_BACSU STANDARD; PRT; 496 AA.  
AC P50729;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ATP-DEPENDENT DNA HELICASE RECO (EC 3.6.1.-) (RECOMBINATION PROTEIN  
DE S).  
GN RECO OR RECO.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / MARRBURG;  
RX MEDLINE=96349105; PubMed=8760912;  
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,  
RA Serror P.;  
RT "Sequence analysis of the Bacillus subtilis chromosome region between  
RT the serA and kds loci cloned in a yeast artificial chromosome.";  
RL Microbiology 142:2005-2016(1996).  
RN [2]





KTRL_YEAST	KTRL_YEAST	STANDARD;	PRT;	393 AA.
AC	P27810;			
DT	01-AUG-1992	(Rel. 23, Created)		
DT	01-AUG-1992	(Rel. 23, Last sequence update)		
DT	20-AUG-2001	(Rel. 40, Last annotation update)		
DE	ALPHA-1,2 MANNOSYLTRANSFERASE KTRL1 (EC 2.4.1.131).			
GN	KTRL1 OR YOR099W OR YOR3189W.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Ascomycota; Saccharomycetaceae; Saccharomycetes.			
OX	NCBI_TaxID=4932;			
RN	[1]			

Db 134 VIPKEHSFP-EWIDEEKAAQVRKEMGEKRIIYGDSISYRHMCRFESGFFYRHPPL---MD 189  
Qy 176 SADMYFMVGHVRV-----IFYVMIDDTSRMPVVLNPLHLSLQVFEIRSEKRWQDISM 226  
Db 190 DYDNYWRVDPDKLHCDIDYDFKFMKDNKKYAF-----ISIKYEATPTLWE----- 240  
Qy 227 MRMTTIGEHILAH--IOHEVDFLCMDVDQ-----VFQDNFGVETLGQLVAQLQAWY 277  
Db 241 ----TTRKMEAHPELIHENNMNLDVSDOGLSYNLCHFWSNFE-----IAALDLWRS 289  
Qy 278 KASPEKTYERRELSAAYIPEGE-----GDF-YYHAAIFGGTPTH 316  
Db 290 PAYSAYFDYLDREGGFFYERWGDAPVHSIGAALFLDRSEIHHFGDIGYHYVPHSHCPIDT 349  
Qy 317 ILNLTRECFKGILQDKKHDIQAQWHDESHNLKYLFLNKPTKILSPEYCWQYQIG 370  
Db 350 SIRLANKC-----DCDPSKDFTHWSYSCTTKFYNNKLPKPAG-----WQNHIG 393

Search completed: May 10, 2002, 11:10:36  
Job time: 204 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 16:06:15 ; Search time 2776.7 Seconds  
(without alignments)  
11304.716 Million cell updates/sec

Title: US-09-863-475A-3  
Perfect score: 1500  
Sequence: 1 CCTCCCTTGACTCTTC.....CTTGACACTATTCTTAACCA 1500

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vi:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1

AR075307

LOCUS

DEFINITION

AR075307

ACCESSION

AR075307

VERSION

AR075307.1

KEYWORDS

GI:10002057

SOURCE

Unknown.

ORGANISM

Unknown.

REFERENCE

1 (bases 1 to 1500)

AUTHORS

Lowe,J.B.

TITLE

Methods and products for the synthesis of oligosaccharide structures on glycoproteins, glycolipids, or as free molecules, and for the isolation of cloned genetic sequences that determine these structures

JOURNAL

Patent: US 5955347-A 3 21-SEP-1999;

FEATURES

Location/Qualifiers

1..1500

source

/organism="unknown"

BASE COUNT

406 a 349 c 374 g 371 t

ORIGIN

linear PAT 28-AUG-2000

AR075307 Sequence  
AX306447 Sequence  
I34188 Sequence 3  
M26925 Mouse galac  
AX128259 Sequence  
AF297615 Mus muscu  
M85153 M.musculus  
AR066335 Sequence  
BC006810 Mus muscu  
L36535 Sus scrofa  
AR048516 Sequence  
AF297614 Mus muscu  
AX128258 Sequence  
A69346 Sequence 3  
A72023 Sequence 3  
AX306443 Sequence  
S71333 alpha 1,3 g  
L36152 Sus scrofa  
AR119910 Sequence  
AX306445 Sequence  
AR066333 Sequence  
AF221509 Sus scrof  
AF221508 Sus scrof  
A69344 Sequence 1  
A72021 Sequence 1  
AX306439 Sequence  
AX306441 Sequence  
J04989 Bovine alph  
AR066334 Sequence  
A69348 Sequence 5  
A72025 Sequence 5  
A69350 Sequence 7  
A72027 Sequence 7  
AR048515 Sequence  
AC012518 Homo sapi  
AX306449 Sequence  
M60263 Human alpha  
AF197007 Pan trogl  
AF196986 Homo sapi  
AX128247 Sequence  
AF221517 Sus scrof  
AX128266 Sequence  
AX128267 Sequence  
AC078879 Homo sapi  
AX128260 Sequence

Query Match 100.0%; Score 1500; DB 6; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 CTTTCCCTTTGAGACTCTTTTGAATGAGAGTACCGATTCTGCTGAAGACCTCGGGCT 60
Db 1 CTTTCCCTTTGAGACTCTTTTGAATGAGAGTACCGATTCTGCTGAAGACCTCGGGCT 60

Qy 61 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCTCTGCTGAGCCCTGCTCCTT 120
Db 61 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCTCTGCTGAGCCCTGCTCCTT 120

Qy 121 AGGAGCCGACAGCTCGACAGAACTCGGTTGCTTTGCTTTGCTTTGGAGGGAACACAG 180
Db 121 AGGAGCCGACAGCTCGACAGAACTCGGTTGCTTTGCTTTGCTTTGGAGGGAACACAG 180

Qy 181 CTGACGATGAGGCTGACTTTGAACCTCAAGAGATCTGTTACCCGAGTCTCGTGAATTA 240
Db 181 CTGACGATGAGGCTGACTTTGAACCTCAAGAGATCTGTTACCCGAGTCTCGTGAATTA 240

Qy 241 AGGCTGTACTACATTTGCTTGGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300
Db 241 AGGCTGTACTACATTTGCTTGGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300

Qy 301 CATCTCAACAAGATCTCCATGTCAAGATCCAAAGTCAGAAACAACTCTTCCATCCTCAAGA 360
Db 301 CATCTCAACAAGATCTCCATGTCAAGATCCAAAGTCAGAAACAACTCTTCCATCCTCAAGA 360

Qy 361 TCTGGATCACAGGAGAAATTAAGTATGTCNAAGGAAAGTAATCCTGTGATGCTGATT 420
Db 361 TCTGGATCACAGGAGAAATTAAGTATGTCNAAGGAAAGTAATCCTGTGATGCTGATT 420

Qy 421 GTCTCAACCGTGTGCTGCTTTTGGGAATATCTCAACAGAATTCACAGAGTTGGTGAG 480
Db 421 GTCTCAACCGTGTGCTGCTTTTGGGAATATGTCNAACAGAATTCACAGAGTTGGTGAG 480

Qy 481 AACAGATGGCAGAAAGGACTGGTGGTCCCAAGCTGGTTTAAAAATGGGACCCACAGTTAT 540
Db 481 AACAGATGGCAGAAAGGACTGGTGGTCCCAAGCTGGTTTAAAAATGGGACCCACAGTTAT 540

Qy 541 CAAGAACAACACGTAGAAAGCAGAGAGAAAGGATAGAAATGGAGATCGCATTTGAAGAG 600
Db 541 CAAGAACAACACGTAGAAAGCAGAGAGAAAGGATAGAAATGGAGATCGCATTTGAAGAG 600

Qy 601 CCTCAGCTATGGGACTGTTTCAATCCAAAGAACCCCGCGGATGTTTACAGTGCACCCCG 660
Db 601 CCTCAGCTATGGGACTGTTTCAATCCAAAGAACCCCGCGGATGTTTACAGTGCACCCCG 660

Qy 661 TGGAGGCGCGGATTTGTTGGGAAGGACATTTATGACACAGCTCTGCTGGGAAAGTACTAC 720
Db 661 TGGAGGCGCGGATTTGTTGGGAAGGACATTTATGACACAGCTCTGCTGGGAAAGTACTAC 720

Qy 721 GCCACAGAAACACTGCTGGGCTGACAGTGTGCTGTGGGAAAGTACATTTGAGCAT 780
Db 721 GCCACAGAAACACTGCTGGGCTGACAGTGTGCTGTGGGAAAGTACATTTGAGCAT 780

Qy 781 TACTTAGAAGACTTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTCATA 840
Db 781 TACTTAGAAGACTTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTCATA 840

Qy 841 TTTTACGTATGATGACGACACCTCCCGGATGCTGTGCTGACCTGAAACCCCTCTACAT 900
Db 841 TTTTACGTATGATGACGACACCTCCCGGATGCTGTGCTGACCTGAAACCCCTCTACAT 900

Qy 901 TCCTTACAAGTCTTTGATGATCAGTCTGTAGAGAGTGGCAGGATATCAGCATATGCGC 960
Db 901 TCCTTACAAGTCTTTGATGATCAGTCTGTAGAGAGTGGCAGGATATCAGCATATGCGC 960

Qy 961 ATGAAGACCATTTGGGAGCAGCATCTGCGCCACATCCAGCAGAGGTCGACTTCTCTTC 1020
Db 961 ATGAAGACCATTTGGGAGCAGCATCTGCGCCACATCCAGCAGAGGTCGACTTCTCTTC 1020
```

```
Qy 1021 TGCATGACGCTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAAACTCTTCGGCCAGCTG 1080
Db 1021 TGCATGACGCTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAAACTCTTCGGCCAGCTG 1080

Qy 1081 GTACACAGCTCCAGGCTGTGTGTACAGGCCAGTCCCGAGAAGTTCCACCTATGAGAG 1140
Db 1081 GTACACAGCTCCAGGCTGTGTGTACAGGCCAGTCCCGAGAAGTTCCACCTATGAGAG 1140

Qy 1141 CGGAACCTGTCGGCGCGGTACATTTCCATTTCGAGAGGGGATTTTACTACACGCGGCC 1200
Db 1141 CGGAACCTGTCGGCGCGGTACATTTCCATTTCGAGAGGGGATTTTACTACACGCGGCC 1200

Qy 1201 ATTTTGGAGAACCCCTTACTACATTTCTCAACCTCAACGAGGAGTGTCTTAAAGGGATC 1260
Db 1201 ATTTTGGAGAACCCCTTACTACATTTCTCAACCTCAACGAGGAGTGTCTTAAAGGGATC 1260

Qy 1261 CTCCAGCAACAACATGACATAGAACCCAGTGGCATGATGAGAGCCACTCAACAAA 1320
Db 1261 CTCCAGCAACAACATGACATAGAACCCAGTGGCATGATGAGAGCCACTCAACAAA 1320

Qy 1321 TACTTCTCTTTCAACAACCCACTAAATCCTATCTCCAGAGTATTGCTGGGACTATCAG 1380
Db 1321 TACTTCTCTTTCAACAACCCACTAAATCCTATCTCCAGAGTATTGCTGGGACTATCAG 1380

Qy 1381 ATAGCCTGCTTCAGATATTTAAAGTGTCAAGGTAGCTTGGCAGACAAAAGATTAAT 1440
Db 1381 ATAGCCTGCTTCAGATATTTAAAGTGTCAAGGTAGCTTGGCAGACAAAAGATTAAT 1440

Qy 1441 TTGCTTAGAATAATGCTGACTTCAAAATTTGATGGAACTTGACACTATTCTTAACCA 1500
Db 1441 TTGCTTAGAATAATGCTGACTTCAAAATTTGATGGAACTTGACACTATTCTTAACCA 1500
```

RESULT 2  
AX306447 AX306447 1500 bp DNA linear PAT 11-DEC-2001  
LOCUS Sequence 9 from Patent WO0188096.  
DEFINITION AX306447  
ACCESSION AX306447  
VERSION AX306447.1 GI:17645668  
KEYWORDS  
SOURCE Musca sp.  
ORGANISM Musca sp.

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Muscoidea; Muscidae; Musca.  
REFERENCE 1 (sites)  
AUTHORS Denning, C. and Clark, J.  
TITLE Animal tissue for xenotransplantation  
JOURNAL Patent: WO 0188096-A 9 22-NOV-2001;  
Geron Corporation (US)

FEATURES  
source Location/Qualifiers  
1..1500  
/organism="Musca sp."  
/db\_xref="taxon:30058"  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CA019042.1"  
/db\_xref="GI:17645668"  
/translation="MTHMLQDLHVNKLSMSRSKSETSLPSSRSQSQEKIMNVKCKVIL  
LMLIVSTVVVFWVEYVNRIPVEGENRWKQWPFPSFKNGTSTQEDNVGRRKGRN  
GDRIEPQLDWDFNPKRPDLVTWPKAPVWEGTYDTALLEYATQKLTGVLTVF  
AVGYIEHYLEDELESADMYFMVGHRYIFYVMIDDTSRMPVHLNPLHSQVFIRES  
KRWQDSMMRMKTIGHEILAHIQHVEDFLFCMDVDVQDFQNFVETLGQVLAQLAAMW  
YKASPEKFTYERRELSAAYTPFEGGDFYHAAIFGPGTPTHLINLTRECFFGILQDKKII  
DICAQHDESHLNKYLFLFNKPTKILSPYCDYQIGLPSDIKSKYKVAWQTKYENLVN  
NV"

BASE COUNT 406 a 349 c 374 g 371 t  
ORIGIN

Query Match 100.0%; Score 1500; DB 6; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CTTCCCTTGTAGACTCTTCTTGGAAATGAGAACTACCGATTCTGCTGAAGACCTCCGGCT	60
Db	1	CTTCCCTTGTAGACTCTTCTTGGAAATGAGAACTACCGATTCTGCTGAAGACCTCCGGCT	60
QY	61	CTCAGGCTCTGGAGATTGGAACCTGTACCTTCCCTTCCCTGCTGCTGAGCCCTGCCCTCCTT	120
Db	61	CTCAGGCTCTGGAGATTGGAACCTGTACCTTCCCTTCCCTGCTGCTGAGCCCTGCCCTCCTT	120
QY	121	AGGCAGCCAGAGCTCGACAGAACTCGGTTGCTTTCCTGCTTTCCTTGGAGGGAACACAG	180
Db	121	AGGCAGCCAGAGCTCGACAGAACTCGGTTGCTTTCCTGCTTTCCTTGGAGGGAACACAG	180
QY	181	CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGCTCTCCTGGAATTA	240
Db	181	CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGCTCTCCTGGAATTA	240
QY	241	AGGCCTGTACTACATTTGGCTGGACCTAAGATTTTCATGATCACTATGCTTCAAGATCTC	300
Db	241	AGGCCTGTACTACATTTGGCTGGACCTAAGATTTTCATGATCACTATGCTTCAAGATCTC	300
QY	301	CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAACAAAGTCTTCCATCCTCAAGA	360
Db	301	CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAACAAAGTCTTCCATCCTCAAGA	360
QY	361	TCTGATCAACAGGAGAAATTAATGAATGTCAAGGGAAGTAACTCTGTTGATGCTGATT	420
Db	361	TCTGATCAACAGGAGAAATTAATGAATGTCAAGGGAAGTAACTCTGTTGATGCTGATT	420
QY	421	GTCTCAACCGTGTGCTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG	480
Db	421	GTCTCAACCGTGTGCTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG	480
QY	481	AACAGATGCGAAGACACTGGTGGTCCCAAGCTGTTTAAATAATGGACCCACACTTAT	540
Db	481	AACAGATGCGAAGACACTGGTGGTCCCAAGCTGTTTAAATAATGGACCCACACTTAT	540
QY	541	CAAGAAGACAACTAGAGGACGAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAG	600
Db	541	CAAGAAGACAACTAGAGGACGAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAG	600
QY	601	CCTCAGCTATGGACACTGTTCAATCAAGAAGCCCGCCGGATGTTTGACAGTGACCCCG	660
Db	601	CCTCAGCTATGGACACTGTTCAATCAAGAAGCCCGCCGGATGTTTGACAGTGACCCCG	660
QY	661	TGGAAGCGCCGATTTGTTGGGAAGCACTTATGACACAGCTCTGCTGGAAAAAGTACTAC	720
Db	661	TGGAAGCGCCGATTTGTTGGGAAGCACTTATGACACAGCTCTGCTGGAAAAAGTACTAC	720
QY	721	GCCACAGAAACTACTCTGGGGCTGACAGTGTGCTGTGGGAAAGTACATTTGAGCAT	780
Db	721	GCCACAGAAACTACTCTGGGGCTGACAGTGTGCTGTGGGAAAGTACATTTGAGCAT	780
QY	781	TACTTTAGAGACTTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCCGGTTCATA	840
Db	781	TACTTTAGAGACTTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCCGGTTCATA	840
QY	841	TTTTTACGTCTATGATAGACACACTCCCGGATGCTGCTGACCTGAAACCTCTTACAT	900
Db	841	TTTTTACGTCTATGATAGACACACTCCCGGATGCTGCTGACCTGAAACCTCTTACAT	900
QY	901	TCCTTACAAGTCTTTGAGATCAGGCTGTGAGAAGAGTGGCAGGATATCAGCATGATGCCG	960
Db	901	TCCTTACAAGTCTTTGAGATCAGGCTGTGAGAAGAGTGGCAGGATATCAGCATGATGCCG	960
QY	961	ATGAAGACCATTTGGGAGCACATCTTGGCCACATCCAGACAGAGTTCGACTTTCCTTTC	1020
Db	961	ATGAAGACCATTTGGGAGCACATCTTGGCCACATCCAGACAGAGTTCGACTTTCCTTTC	1020
QY	1021	TGCATGGAGTGCATCAAGTCTTTCAAGACAACTTGGGGTGGAAACTCTGGGCCAGCTG	1080
Db	1021	TGCATGGAGTGCATCAAGTCTTTCAAGACAACTTGGGGTGGAAACTCTGGGCCAGCTG	1080

QY	1081	GTAGCAGCTCCAGGCTCGTGGTACAAAGCCAGTCCCGAGAAGTTCACCTATGAGAG	1140
Db	1081	GTAGCAGCTCCAGGCTCGTGGTACAAAGCCAGTCCCGAGAAGTTCACCTATGAGAG	1140
QY	1141	CGGAACTCTGCGCCCGGTACATTTCCATTCGGAGAGGGGATTTTACTACACGCGGCC	1200
Db	1141	CGGAACTCTGCGCCCGGTACATTTCCATTCGGAGAGGGGATTTTACTACACGCGGCC	1200
QY	1201	ATTTTGGAGGAACGCTACTACATTTCAACCTCACCAGGAGTGCTTTAAGGGGATC	1260
Db	1201	ATTTTGGAGGAACGCTACTACATTTCAACCTCACCAGGAGTGCTTTAAGGGGATC	1260
QY	1261	CTCCAGGACAAGAACATGACATAGAACCCAGTGGCATGATGAGAGCCACCTCAACAA	1320
Db	1261	CTCCAGGACAAGAACATGACATAGAACCCAGTGGCATGATGAGAGCCACCTCAACAA	1320
QY	1321	TACTTCTTTTCAACAAACCCACTAAATCCTATCTCCAGAGTATTTGCTGGGACTATCAG	1380
Db	1321	TACTTCTTTTCAACAAACCCACTAAATCCTATCTCCAGAGTATTTGCTGGGACTATCAG	1380
QY	1381	ATAGGCTCGCTTCAGATATTAAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAAT	1440
Db	1381	ATAGGCTCGCTTCAGATATTAAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAAT	1440
QY	1441	TTGGTTAGAAATAATGCTGACTTCAAAATGTGATGGAACCTTGACACTATTTCTAACCA	1500
Db	1441	TTGGTTAGAAATAATGCTGACTTCAAAATGTGATGGAACCTTGACACTATTTCTAACCA	1500
RESULT	3		
LOCUS	I34188	Sequence 3 from patent US 5595900.	DNA 1linear PAT 06-FEB-1997
DEFINITION	I34188		
ACCESSION	I34188		
VERSION	I34188.1	GI:1824979	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1500)		
AUTHORS	Lowe,J.B.		
TITLE	Methods and products for the synthesis of oligosaccharide structures on glycoproteins, glycolipids, or as free molecules, and for the isolation of cloned genetic sequences that determine these structures		
JOURNAL	Patent: US 5595900-A 3 21-JAN-1997;		
FEATURES	source	Location/Qualifiers	
	1..1500	/organism="unknown"	
BASE COUNT	406 a 349 c 374 g 371 t		
ORIGIN			
Query Match	100.0%;	Score 1500;	DB 6; Length 1500;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1500;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	CTTTCCTTGTAGACTCTTCTTGGAAATGAGAACTACCGATTCTGCTGAAGACCTCCGGCT	60
Db	1	CTTTCCTTGTAGACTCTTCTTGGAAATGAGAACTACCGATTCTGCTGAAGACCTCCGGCT	60
QY	61	CTCAGGCTCTGGAGATTGGAACCTGTACCTTCCCTTCCCTGCTGCTGAGCCCTGCCCTCCTT	120
Db	61	CTCAGGCTCTGGAGATTGGAACCTGTACCTTCCCTTCCCTGCTGCTGAGCCCTGCCCTCCTT	120
QY	121	AGGCAGCCAGAGCTCGACAGAACTCGGTTGCTTTCCTGCTTTCCTTGGAGGGAACACAG	180
Db	121	AGGCAGCCAGAGCTCGACAGAACTCGGTTGCTTTCCTGCTTTCCTTGGAGGGAACACAG	180
QY	181	CTGACCATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGCTCTCCTGGAATTA	240
Db	181	CTGACCATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGCTCTCCTGGAATTA	240

```
Qy 241 AGGCGTGACTACATTTGCCCTGGACCTAAGATTTTTCATGATCACTATGCTTCAAGATCTC 300
Db 241 AGGCGTGACTACATTTGCCCTGGACCTAAGATTTTTCATGATCACTATGCTTCAAGATCTC 300
Qy 301 CATGTCACACAGATCTCCATGTCGAAGTCCCAAGTCACAGAACAGCTCTCCATCCTCAAGA 360
Db 301 CATGTCACACAGATCTCCATGTCGAAGTCCCAAGTCACAGAACAGCTCTCCATCCTCAAGA 360
Qy 361 TCTGGATCACAGGAGAAAATTAATGAATGTCGAAGGAAAAGTAATCCTGTTGATGCTGATT 420
Db 361 TCTGGATCACAGGAGAAAATTAATGAATGTCGAAGGAAAAGTAATCCTGTTGATGCTGATT 420
Qy 421 GTCTCAACCGTGTGCTGTTTGGGAATATGTCACAGAAATTCACAGAGTTGGTGAG 480
Db 421 GTCTCAACCGTGTGCTGTTTGGGAATATGTCACAGAAATTCACAGAGTTGGTGAG 480
Qy 481 AACAGATGGCAGAGGACTGCTGTTCCCAAGCTGGTTTAAANAATGGACCCACAGTTAT 540
Db 481 AACAGATGGCAGAGGACTGCTGTTCCCAAGCTGGTTTAAANAATGGACCCACAGTTAT 540
Qy 541 CAAGAACAACGCTAGAAGACGAGAGAAAAGGTAGAAATGGAGATCGCATTTGAAGAG 600
Db 541 CAAGAACAACGCTAGAAGACGAGAGAAAAGGTAGAAATGGAGATCGCATTTGAAGAG 600
Qy 601 CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGCGGATGTTTTCAGAGTGACCCCG 660
Db 601 CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGCGGATGTTTTCAGAGTGACCCCG 660
Qy 661 TGGAGGCGCGGATGCTGGGAGGACCTTATGACACAGCTCTGCTGGAAAAGTACTAC 720
Db 661 TGGAGGCGCGGATGCTGGGAGGACCTTATGACACAGCTCTGCTGGAAAAGTACTAC 720
Qy 721 GCCACACAGAACTCACTGTGGGCTGACAGTGTTGCTGTGGAAAAGTACATTTGAGCAT 780
Db 721 GCCACACAGAACTCACTGTGGGCTGACAGTGTTGCTGTGGAAAAGTACATTTGAGCAT 780
Qy 781 TACTTAGAAGACTTTCGAGCTGCTGACATGATGATTTTTCATGTTGGCCATCGGGTCATA 840
Db 781 TACTTAGAAGACTTTCGAGCTGCTGACATGATGATTTTTCATGTTGGCCATCGGGTCATA 840
Qy 841 TTTTACGTCATGATAGACGACACCTCCCGGATGCTGCTGTCGACCTGAACCCCTCATAT 900
Db 841 TTTTACGTCATGATAGACGACACCTCCCGGATGCTGCTGTCGACCTGAACCCCTCATAT 900
Qy 901 TCCTTACAAGTCTTTGAGATCAGTCTGAGAAAGAGTGGCAGGATATCAGCATATGCGC 960
Db 901 TCCTTACAAGTCTTTGAGATCAGTCTGAGAAAGAGTGGCAGGATATCAGCATATGCGC 960
Qy 961 ATGAAGACCATTTGGGAGCACATCCTGGCCACATCCAGACAGAGTTCGACTTCTCTTC 1020
Db 961 ATGAAGACCATTTGGGAGCACATCCTGGCCACATCCAGACAGAGTTCGACTTCTCTTC 1020
Qy 1021 TGCATGGACGTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAACCTCTGGGCCACGCTG 1080
Db 1021 TGCATGGACGTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAACCTCTGGGCCACGCTG 1080
Qy 1081 GTAGCACAGCTCCAGGCTGGTGGTACAAGGCCAGTCCCGAGAGTTTCACTATGAGAGG 1140
Db 1081 GTAGCACAGCTCCAGGCTGGTGGTACAAGGCCAGTCCCGAGAGTTTCACTATGAGAGG 1140
Qy 1141 CGGGAACGTGGCGCGGTACATTTCCATTCGGAGAGGGGATTTTACTACCGCGGCC 1200
Db 1141 CGGGAACGTGGCGCGGTACATTTCCATTCGGAGAGGGGATTTTACTACCGCGGCC 1200
Qy 1201 ATTTTTCGAGGAACGCTACTACATTTCTCAACCTCACAGGGAGTCTTTAAGGGGATC 1260
Db 1201 ATTTTTCGAGGAACGCTACTACATTTCTCAACCTCACAGGGAGTCTTTAAGGGGATC 1260
Qy 1261 CTCCAGGACAAAGAACATGACATAGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAA 1320
Db 1261 CTCCAGGACAAAGAACATGACATAGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAA 1320
```

```
Qy 1321 TACTTCTTTTCAACAAACCCACTAAAATCCTATCTCCAGAGTATTGCTGGGACTATCAG 1380
Db 1321 TACTTCTTTTCAACAAACCCACTAAAATCCTATCTCCAGAGTATTGCTGGGACTATCAG 1380
Qy 1381 ATAGCCCTGCCCTTCAGATATTAAAAGTCTCAAGGTAGCTTGGCAGACAAAAGATATAAT 1440
Db 1381 ATAGCCCTGCCCTTCAGATATTAAAAGTCTCAAGGTAGCTTGGCAGACAAAAGATATAAT 1440
Qy 1441 TTGGTTAGAAATAATGCTGACTTCAAAATTTGATGGAACCTTGACACTATTTCTTAACCA 1500
Db 1441 TTGGTTAGAAATAATGCTGACTTCAAAATTTGATGGAACCTTGACACTATTTCTTAACCA 1500

RESULT 4
MUSGALA
LOCUS MUSGALA 1500 bp mRNA linear ROD 12-JUN-1993
DEFINITION Mouse galactosyltransferase mRNA, complete cds.
ACCESSION M26925
VERSION M26925.1 GI:193419
KEYWORDS galactosyl transferase.
SOURCE Mouse F9 teratocarcinoma, cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Larsen,R.D., Rajan,V.P., Ruff,M.M., Kukowaka-Latallo,J.,
Cummings,R.D. and Lowe,J.B.
TITLE Isolation of a cDNA encoding a murine UDPgalactose:beta-D-
galactosyl-1,4-N-acetyl-D-glucosaminide alpha-1,3-
galactosyltransferase: Expression cloning by gene transfer
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86, 8227-8231 (1989)
MEDLINE 90046769
COMMENT Draft entry and computer-readable copy of sequence [1] kindly
submitted by J.B.Lowe, 08-AUG-1989.
FEATURES
source Location/Qualifiers
1..1500
/organism="Mus musculus"
/db_xref="taxon:10090"
277..1461
/notes="galactosyltransferase (EC 2.4.1.151)"
/codon_start=1
/protein_id="AA37657.1"
/db_xref="GI:309242"
/translation="MITLQDLHVNKISMSRSKSETSLPSRSQSQEKINVKVKVIL
LMILIVSTVVVFWYVYRNPVEVGNKQWPFNSFKNGTHSYEDNVGKREKGRN
GRIEELPQLWDFNPKRPDLVTPMKQPIVWEGTYDTALKEYYATKLTGLTVF
AVGKYIEHYLEDFLESADMYFMVGHVIFVYMDVTSRMPVHLNPLHLSQVPIRSE
KQKQDISMRMKTIGEHILAHIOHEVDVFLFMDVDVQFQDFGVETIGQLVAOLOAAM
YKASPEKFTYERBELSAAYTPFGEGDYYHAATFEGTPTHLNLITRECFKGILODKKH
DIEAQHDESHLNKYFLFNKPTKILSPYCWIDYQIGLPSDIKSVKVAQWNTREYNLVRN
NV"
BASE COUNT 406 a 349 c 374 g 371 t
ORIGIN
Query Match 100.0%; Score 1500; DB 10; Length 1500;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCTTCCCTCTGAGACTCTCTTGGAAATCAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60
Db 1 CCTTCCCTCTGAGACTCTCTTGGAAATCAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60
Qy 61 CTCAGGCTCTGGGAGTTTGGAAACCTGTACCTTCCTTCTCTGCTGAGCCCTGCTCCTT 120
Db 61 CTCAGGCTCTGGGAGTTTGGAAACCTGTACCTTCCTTCTCTGCTGAGCCCTGCTCCTT 120
Qy 121 AGGAGGCCAGAGCTCGACAGAACTCGGTTGCTTGTGTTGCTTTGGAGGGGAACACAG 180
Db 121 AGGAGGCCAGAGCTCGACAGAACTCGGTTGCTTGTGTTGCTTTGGAGGGGAACACAG 180
Qy 181 CTGACGATGAGGCTGACTTTTGAACCTCAAGAGATCTGCTTACCCAGTCTCCTCGAATTA 240
Db 181 CTGACGATGAGGCTGACTTTTGAACCTCAAGAGATCTGCTTACCCAGTCTCCTCGAATTA 240
```

[illegible]

Query Match	95.7%	Score 1434.8	DB 6	Length 3537
Best Local Similarity	97.5%	Pred. No. 0		
Matches 1493	Conservative 0	Mismatches 2	Indels 37	Gaps 2
Qy 1	CTTTCCCTTGTAGACTCTTCTTGGNAATGAGAAGTACCGATTTCTGCTGAAGACCTCGCGCT	60		
Db				
Qy 120	CTTTCCCTTGTAGACTCTTCTTGGNAATGAGAAGTACCGATTTCTGCTGAAGACCTCGCGCT	179		
Db				
Qy 61	CTCAGGCTCTGGGAGTGTGGAACCCGTGTAACCTTCCTTTCCTGCTGAGCCCTGCTCCCTT	120		
Db				
Qy 180	CTCAGGCTCTGGGAGTGTGGAACCCGTGTAACCTTCCTTTCCTGCTGAGCCCTGCTCCCTT	239		
Db				
Qy 121	AGGCAGGCCAGAGCTGCACAGAA-CTCGGTTGCTTTGCTGTTGCTTTGGAGGGAACACA	179		
Db				
Qy 240	AGGCAGGCCAGAGCTGCAGAGAACTCCGGTTGCTTTGCTTTGCTTTGGAGGGAACACA	299		
Db				
Qy 180	GCTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGATCTCCTGGAATTA	239		
Db				
Qy 300	GCTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGATCTCCTGGAATTA	359		
Db				
Qy 240	AAGGCTCTACTACATTTTGCCTGGACCTTAAGATTTTTCATGATCATGCTTCAAGATCT	299		
Db				
Qy 360	AAGGCTCTACTACATTTTGCCTGGACCTTAAGATTTTTCATGATCATGCTTCAAGATCT	419		
Db				
Qy 300	CCATGTCACAGATCTCCATGTCGAAGTCCAAGTCAGAAACAAGCTTTCCATCTCCAAG	359		
Db				
Qy 420	CCATGTCACAGATCTCCATGTCGAAGTCCAAGTCAGAAACAAGCTTTCCATCTCCAAG	479		
Db				
Qy 360	ATCTCGATCACAGGAAATAAATCAATGCTCAAGGAAAGCTTAATCCCTGTTGATGCTGAT	419		
Db				
Qy 480	ATCTCGATCACAGGAAATAAATCAATGCTCAAGGAAAGCTTAATCCCTGTTGATGCTGAT	539		
Db				
Qy 420	TGCTTCAACCGTGTGTCGTTGTTGGGAATATGCTCAACAG-----461			
Db				

```
Db 540 TGTCTCAACCGTGGTTCGTCGTTTTCGGGAATATGTCAACAGCCCAACGCGCTCTTCTTT 599
Qy 462 -----AATTCAGAGTTGGTGTGAGACAGATGGCAGAGGACTGGTG 503
Db 600 GTGGATATATACACAAAAATTCAGAGTTGGTGTGAGACAGATGGCAGAGGACTGGTG 659
Qy 504 GTTCCCAAGCTGTTTAAAAATGGGACCCACAGTTTCAAGAGACACACGTAGAACGACG 563
Db 660 GTTCCCAAGCTGTTTAAAAATGGGACCCACAGTTTCAAGAGACACACGTAGAACGACG 719
Qy 564 GAGAGAAAGGTTAGAAATCGAGATCCGATTGAAGACGCTCAGCTATGGGACTGGTTCAA 623
Db 720 GAGAGAAAGGTTAGAAATCGAGATCCGATTGAAGACGCTCAGCTATGGGACTGGTTCAA 779
Qy 624 TCCAAAGAACCGCCCGGATGTTTGGACAGTGACCCCGTGAAGGCGCGGATTTGTGGGA 683
Db 780 TCCAAAGAACCGCCCGGATGTTTGGACAGTGACCCCGTGAAGGCGCGGATTTGTGGGA 839
Qy 684 AGGCACATTATGACACAGCTCTGTGGAAAAAGTACTACGCCACACAGAAATCACTGTGGG 743
Db 840 AGGCACATTATGACACAGCTCTGTGGAAAAAGTACTACGCCACACAGAAATCACTGTGGG 899
Qy 744 GCTGACAGTGTCTGCTGGGAAAGTACATTGAGCATTACTTAGAACCTTTCTGGAGTC 803
Db 900 GCTGACAGTGTCTGCTGGGAAAGTACATTGAGCATTACTTAGAACCTTTCTGGAGTC 959
Qy 804 TGCTGACATGTACTTTCATGTTGGCCATCGGGTCATATTTTACGTCATGATAGACGAC 863
Db 960 TGCTGACATGTACTTTCATGTTGGCCATCGGGTCATATTTTACGTCATGATAGACGAC 1019
Qy 864 CTCCCGGATGCTGTGTCGACCTGAACCTCTACATTCCTTACAACTCTTTGAGATCAG 923
Db 1020 CTCCCGGATGCTGTGTCGACCTGAACCTCTACATTCCTTACAACTCTTTGAGATCAG 1079
Qy 924 GTCTGAGAAAGGTGGCAGATATCAGCATGATCGCATGAAGACCATGGGGAGCACAT 983
Db 1080 GTCTGAGAAAGGTGGCAGATATCAGCATGATCGCATGAAGACCATGGGGAGCACAT 1139
Qy 984 CTTGGCCACATCCAGCAGAGTTCGACTTCTCTTCTGTCATGGACGCTGATCAAGTCTT 1043
Db 1140 CTTGGCCACATCCAGCAGAGTTCGACTTCTCTTCTGTCATGGACGCTGATCAAGTCTT 1199
Qy 1044 TCAAGACAACTTCGGGGTGAAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCGCTGGT 1103
Db 1200 TCAAGACAACTTCGGGGTGAAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCGCTGGT 1259
Qy 1104 GTACAAGCCAGTCCCGAGAGTTCACCTATGAGAGCGGGAACCTGTCGGCGCGGTACAT 1163
Db 1260 GTACAAGCCAGTCCCGAGAGTTCACCTATGAGAGCGGGAACCTGTCGGCGCGGTACAT 1319
Qy 1164 TCCATTCCGAGAGGGGATTTTACTACACGCGGCCATTTTGGAGGAACGCGCTACTCA 1223
Db 1320 TCCATTCCGAGAGGGGATTTTACTACACGCGGCCATTTTGGAGGAACGCGCTACTCA 1379
Qy 1224 CATTTCTCAACTCACAGGAGTGCTTTTAAAGGGATCCTCCAGGACAGAAACATGACAT 1283
Db 1380 CATTTCTCAACTCACAGGAGTGCTTTTAAAGGGATCCTCCAGGACAGAAACATGACAT 1439
Qy 1284 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAACCCAC 1343
Db 1440 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAACCCAC 1499
Qy 1344 TAAATCCTATCTCCAGAGATTTGCTGGGACTATCAGATAGGCGCTGCTTTCAGATATTA 1403
Db 1500 TAAATCCTATCTCCAGAGATTTGCTGGGACTATCAGATAGGCGCTGCTTTCAGATATTA 1559
Qy 1404 AAGTGTCAAGGTAGCTTTGGCAGACAAAAGAGTATAATTTGGTTAGAAATATGCTGACT 1463
Db 1560 AAGTGTCAAGGTAGCTTTGGCAGACAAAAGAGTATAATTTGGTTAGAAATATGCTGACT 1619
Qy 1464 TCAAAATTTGATGGAAACTTTGACACTATTTCT 1495
Db 1620 TCAAAATTTGATGGAAACTTTGACACTATTTACT 1651
```

## RESULT 6

```
AF297615 LOCUS
DEFINITION Mus musculus alpha-1,3-galactosyltransferase gene, complete cds.
ACCESSION AF297615
VERSION AF297615.1 GI:15419873
KEYWORDS house mouse.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3537)
AUTHORS Koike,C., Friday,R., Fung,J.J., Starzl,T.E. and Trucco,M.
TITLE Comparison of the regulatory regions of the
alpha1,3galactosyltransferase gene between murine and porcine
species
JOURNAL Transplant. Proc. 33 (1-2), 710-711 (2001)
MEDLINE 21168575
PUBMED 11267031
REFERENCE 2 (bases 1 to 3537)
AUTHORS Koike,C., Starzl,T.E. and Trucco,M.
TITLE Isolation and characterization of the regulatory region of the
murine alpha-1,3-galactosyltransferase gene
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3537)
AUTHORS Koike,C., Starzl,T.E. and Trucco,M.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-2000) Surgery, University of Pittsburgh, 3601
Fifth Ave., Pittsburgh, PA 15215, USA
FEATURES
source
1..3537
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
502..1617
/codon_start=1
/product="alpha-1,3-galactosyltransferase"
/protein_id="AAK97206.1"
/db_xref="GI:15419874"
/translation="MNVGKVIILLIIVTVVVFVEYVNSPDGSEFLWIYHTKIPEVG
ENRWKQDMFPSPFNGTNGTISQEDNVGREGKRNDRLEEPOLWDFNFKNRPDLVT
VTPKAPIWEGTIDTALKEYYATOKLTGLTVFVAGKYIEHYLEDFLDSADMYFMV
GRHVFYVMDITSRMPVHLNPLHLSQVFEIRSEKRWODISMHRKTIIGEHTLAHIQ
HEYDFLCMDVDQVFNFGVETLGLOLAQWAKASPEKTYRRELISAAYIPFG
EGDFYVHAATFGCTPHTILNLTRECFKGLQDKKHDEAQMHDESHLNKFLFNKPTK
ILSPEYCDYQIGLPSDIKSVKAWQTKYENLVYRNW"
polyA_signal 3419..3424
polyA_site 3521..3527
BASE COUNT 978 a 777 c 843 g 939 t
ORIGIN
```

```
Query Match 95.7%; Score 1434.8; DB 10; Length 3537;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1493; Conservative 0; Mismatches 2; Indels 37; Gaps 2;

Qy 1 CCTTCCCTCTGAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60
Db 120 CCTTCCCTCTGAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 179
Qy 61 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCCTTCTCTGCTGAGCCCTGCCTCCTT 120
Db 180 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCCTTCTCTGCTGAGCCCTGCCTCCTT 239
Qy 121 AGGAGCCGACAGCTGCACAGAA-CTCGTGTCTTTGCTGTGTTTCTTTGGAGGAAACACA 179
Db 240 AGGAGCCGACAGCTGCACAGAACTCGTGTCTTTGCTGTGTTTCTTTGGAGGAAACACA 299
Qy 180 GCTGACCATGAGGCTGACTTTGAACCTCAAGAGATCTGCTTACCCAGTCTCCTCGAATTA 239
```

Db 300 GCTGACGATGAGGCTGACTTTTGAACCTCAAGAGATCTGCTTACCCAGTCTCCTGGAAATTA 359  
QY 240 AAGGCTGTACTACATTTGCTGGACCTAAGATTTTCATGATCACTATGCTTCAAGATCT 299  
Db 360 AAGGCTGTACTACATTTGCTGGACCTAAGATTTTCATGATCACTATGCTTCAAGATCT 419  
QY 300 CCATGTCAACAAGATCTCCATGTCAAGATCCAAAGTCAAGTCAAGTCAAGTCAAGTCAAG 359  
Db 420 CCATGTCAACAAGATCTCCATGTCAAGATCCAAAGTCAAGTCAAGTCAAGTCAAGTCAAG 479  
QY 360 ATCTGATCACAGAGAAATTAATGAATGTCAAGGAAAGATTAATCTGTGTGATGCTGAT 419  
Db 480 ATCTGATCACAGAGAAATTAATGAATGTCAAGGAAAGATTAATCTGTGTGATGCTGAT 539  
QY 420 TGCTCAACCGTGGTGTGCTGTTTTGGGAATATGTCAACAG----- 461  
Db 540 TGCTCAACCGTGGTGTGCTGTTTTGGGAATATGTCAACAGCCAGACGGCTCTTCTT 599  
QY 462 -----AATTCAGAGTGTGTTGAGACAGATGGCAGAAAGACTGGTG 503  
Db 600 GTGGATATATCACACAAAATTCAGAGGTTGGTGAACAGATGGCAGAAAGACTGGTG 659  
QY 504 GTTCCCAAGCTGGTTAAATAATGGACCCACACTTATCAAGAGACAAAGTCAAGAGCAG 563  
Db 660 GTTCCCAAGCTGGTTAAATAATGGACCCACACTTATCAAGAGACAAAGTCAAGAGCAG 719  
QY 564 GAGAGAAAAGGTAGAAATGGAGATCGCATTTGAAGAGCCCTCAGCTATGGGACTGGTTCAA 623  
Db 720 GACAGAAAAGGTAGAAATGGAGATCGCATTTGAAGAGCCCTCAGCTATGGGACTGGTTCAA 779  
QY 624 TCCAAAGAACCGCCGGATGTTTTGACAGTGACCCCGTGAAGCGCCGATTTGTGGGA 683  
Db 780 TCCAAAGAACCGCCGGATGTTTTGACAGTGACCCCGTGAAGCGCCGATTTGTGGGA 839  
QY 684 AGGCACTTATGACACAGCTCTGCTGGAAGAGTACTAGCCACACAGAACTCACTGTGG 743  
Db 840 AGGCACTTATGACACAGCTCTGCTGGAAGAGTACTAGCCACACAGAACTCACTGTGG 899  
QY 744 GCTGACAGTGTGCTGTTGGGAAAGTACATTTAGACATTTACTTGAAGACTTTCTGGAGTC 803  
Db 900 GCTGACAGTGTGCTGTTGGGAAAGTACATTTAGACATTTACTTGAAGACTTTCTGGAGTC 959  
QY 804 TGCTGACATGTACTTCAATGTTGGCCATCGGGTCAATTTTACGTCATGATGACGACAC 863  
Db 960 TGCTGACATGTACTTCAATGTTGGCCATCGGGTCAATTTTACGTCATGATGACGACAC 1019  
QY 864 CTCCGGATGCTGCTGACCTGAACCTCTACATTTCCCTTACAAGTCTTTTCAGATCAG 923  
Db 1020 CTCCGGATGCTGCTGACCTGAACCTCTACATTTCCCTTACAAGTCTTTTCAGATCAG 1079  
QY 924 GTCTGAGAGAGGTGGCAGGATATCAGCATGATCGGCATGAAGACCATTTGGGGAGCAGAT 983  
Db 1080 GTCTGAGAGAGGTGGCAGGATATCAGCATGATCGGCATGAAGACCATTTGGGGAGCAGAT 1139  
QY 984 CTTGGCCACATCCAGACAGGTGCTGCTTCTTCTGATGGAGTGGATCAAGTCTTT 1043  
Db 1140 CTTGGCCACATCCAGACAGGTGCTGCTTCTTCTGATGGAGTGGATCAAGTCTTT 1199  
QY 1044 TCAAGCAACTCTCGGGTGGAACTCTGGGCCAGCTGGTAGCAGCTCCAGGCTGGTG 1103  
Db 1200 TCAAGCAACTCTCGGGTGGAACTCTGGGCCAGCTGGTAGCAGCTCCAGGCTGGTG 1259  
QY 1104 GTAAAGGCGAGTCCCGAGAAAGTTACCTATGAGAGCGGGAAGTCTGCGGCCGCTACAT 1163  
Db 1260 GTAAAGGCGAGTCCCGAGAAAGTTACCTATGAGAGCGGGAAGTCTGCGGCCGCTACAT 1319  
QY 1164 TCCATTGGAGAGGGGATTTTACTACCACGCGGCCATTTTGGAGGAACGCTACTCA 1223  
Db 1320 TCCATTGGAGAGGGGATTTTACTACCACGCGGCCATTTTGGAGGAACGCTACTCA 1379  
QY 1224 CATTTCAACCTCACCAGGGAGTGTCTTTAAGGGATCTCCAGGACAGAAACATGACAT 1283  
Db 1380 CATTTCAACCTCACCAGGGAGTGTCTTTAAGGGATCTCCAGGACAGAAACATGACAT 1439

QY 1284 AGAAGCCAGTGGCATGATGAGCCACCTCAACAAATACTTCTCTTTCAACAAACCCAC 1343  
Db 1440 AGAAGCCAGTGGCATGATGAGCCACCTCAACAAATACTTCTCTTTCAACAAACCCAC 1499  
QY 1344 TAAATCTCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCGCTTCAGATATAA 1403  
Db 1500 TAAATCTCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCGCTTCAGATATAA 1559  
QY 1404 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAATTTGTTAGAAATATGCTGACT 1463  
Db 1560 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAATTTGTTAGAAATATGCTGACT 1619  
QY 1464 TCAATTTGTGATGGAACCTTGACACTATTCTT 1495  
Db 1620 TCAATTTGTGATGGAACCTTGACACTATTACT 1651  
RESULT 7  
MUSGLYTNLS  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
FEATURES  
Source  
1. .3450  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/cell\_line="C127"  
1. .3450  
/gene="Gata-1"  
445. .1560  
/gene="Gata-1"  
/note="positions 463-495 code for the transmembrane domain"  
/codon\_start=1  
/product="alpha-1,3-galactosyltransferase"  
/protein\_id="AA37711.1"  
/translation="MNVKGVILLIVTVVVFWEYVNSPDGSLWLYHTKIPVVG  
ENRQKDNHFPSPKNGTHSYQEDNVEGRKRCNDRIEERPOLMDFNPKRPDVL  
VTPWKAPVWEGYDTALLKLYATOKLVGLTVFVAGKYIEHYLEDPLSDMYFMV  
GHRVIFYMDDVTSRPMVHLNPLSLQVFEIRSEKQWDISMRMKTIGEHLAHIQ  
HEVDFLCMDVQFDNFVETLGQVLAQAWKYKASPKFTYRRRLSAAYIPFG  
EDDFYHAAIFGFTPTTHILNRECFKGILQDKHIDIEAQWHDESHLNYFLFNKPTK  
ILSPEYCDWYQIGLPSDIKSVKVAWOTKEYNLVRNV"  
BASE COUNT 954 a 762 c 826 g 908 t  
ORIGIN chromosome 2 (centromeric region).  
Query Match 94.6%; Score 1419.6; DB 10; Length 3450;  
Best Local Similarity 97.3%; Pred. No. 0;  
Matches 1490; Conservative 0; Mismatches 4; Indels 38; Gaps 3;  
QY 1 CCTTCCCTTGTAGACTCTTCTTGAATGAGAAGTACCAGTATCTCTGCTGAAGACCTCGCGCT 60  
Db 64 CCTTCCCTTGTAGACTCTTCTTGAATGAGAAGTACCAGTATCTGCTGAAGACCTCGCGCT 123  
QY 61 CTGAGGCTCTGGAGATTGGAAACCTGTACCTTCTCTTCTGCTGAGCCCTGCTCTCTT 120  
Db 124 CTCAGGCTCTGGAGATTGGAAACCTGTACCTTCTCTTCTGCTGAGCCCTGCTCTCTT 183



Qy	121	AGGCAGGCCAGAGCTCCACAGAA - CTCGGTTGCTTTGCTGTTTGCCTTTGTTGGAGGAACACA	179
Db	184	CGGCAGGCCAGAGCTCCACAGAAGCTCGGTTGCTTTGCTGTTTGCCTTTGGAGGAACACA	243
	180	GCTGACGATGAGGCTGACTTTGAATCAAGAGATCTGCTTACCCCAAGTCCTCTGGAATTA	239
Db	244	GCTGACGATGAGGCTGACTTTGAATCAAGAGATCTGCTTACCCCAAGTCCTCTGGAATTA	303
Qy	240	AAGGCTGTACTACATTTGCTGGAGCTAAGATTTTCATGATCACTATGCTTCAAGATCT	299
Db	304	AAGGCTGTACTAC - CTTGCTGGACCTAAGATTTTCATGATCACTATGCTTCAAGATCT	362
Qy	300	CCATGTCACAAGATCTCCATGTCGAAGTCCAAGTCAAGAACAAAGTCCTTCCATCCTCAAG	359
Db	363	CCATGTCACAAGATCTCCATGTCGAAGTCCAAGTCCAAGTCAAGAACAAAGTCCTTCCATCCTCAAG	422
Qy	360	ATCTGGATCACAGGAGAAATATGAATCTCAAGGAAAGTAATCTCTGTTGATGCTGAT	419
Db	423	ATCTGGATCACAGGAGAAATATGAATCTCAAGGAAAGTAATCTCTGTTGATGCTGAT	482
Qy	420	TGCTCAACCGTGGTTGCTGCTGTTTGGGAATATGTCACAG - -----	461
Db	483	TGCTCAACCGTGGTTGCTGCTGTTTGGGAATATGTCACAGCCAGAGCGCTCTTCTTT	542
Qy	462	----- -AATTCAGAGTTGCTGAGAACAGATGGCAGAAGGACTCTGGTG	503
Db	543	GTGGATATATCACACAAAAATTCAGAGGTTTGGTGAGAACAGATGGCAGAAGGACTGGTG	602
Qy	504	GTTCCTCAAGCTGCTTTAAAATGGGACCCACAGTTATCAAGAAGACAACGTAGAAGGACG	563
Db	603	GTTCCTCAAGCTGCTTTAAAATGGGACCCACAGTTATCAAGAAGACAACGTAGAAGGACG	662
Qy	564	GAGAAAAAGGGTAGAATGGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAA	623
Db	663	GAGAAAAAGGGTAGAATGGAGATCGCATTTCAAGAGCCTCAGCTATGGGACTGGTTCAA	722
Qy	624	TCCAAAGAACGCCCGGATGTTTGACAGTGACCCGCTGGAAGGGCGGATTTGTGTGGGA	683
Db	723	TCCAAAGAACGCCCGGATGTTTGACAGTGACCCGCTGGAAGGGCGGATTTGTGTGGGA	782
Qy	684	AGGCACCTATTGACACAGCTCTGCTGGAAAAGTACTACGCCACACAGAAACTCACTGTGGG	743
Db	783	AGGCACCTATTGACACAGCTCTGCTGGAAAAGTACTACGCCACACAGAAACTCACTGTGGG	842
Qy	744	GCTGACAGCTGTTGCTGTGGGAAAGTACATTTAGAGCATTTACTTTAGAAGACTTTCTGGAGTC	803
Db	843	GCTGACAGCTGTTGCTGTGGGAAAGTACATTTAGAGCATTTACTTTAGAAGACTTTCTCTGGAGTC	902
Qy	804	TGCTGACATGCTACTTCATGTTTGGCCATCGGCTCATATTTACGTCATGATAGACGAC	863
Db	903	TGCTGACATGCTACTTCATGTTTGGCCATCGGCTCATATTTACGTCATGATAGACGAC	962
Qy	864	CTCCCGGATGCCTGTGCTGCACCTGAACCCCTACATTTCTTACAAGCTTTTGAGATCAG	923
Db	963	CTCCCGGATGCCTGTGCTGCACCTGAACCCCTACATTTCTTACAAGCTTTTGAGATCAG	1022
Qy	924	GTCTGAGAAGAGTGGCAGGATATACGATATGATGGCATGAAGACCATTTGGGGAGCACAT	983
Db	1023	GTCTGAGAAGAGTGGCAGGATATACGATATGATGGCATGAAGACCATTTGGGGAGCACAT	1082
Qy	984	CCTGGCCACATCCAGACAGGTGACATTCCTCTTCTGCAATGGACGTGGATCAAGTCTT	1043
Db	1083	CCTGGCCACATCCAGACAGGTGACATTCCTCTCTCTGCAATGGACGTGGATCAAGTCTT	1142
Qy	1044	TCAGACAACTTCGGGGTGGAAACTCTGGGCCAGCTGGTAGCACAGCTTCCAGGCGCTGGTG	1103
Db	1143	TCAGACAACTTCGGGGTGGAAACTCTGGGCCAGCTGGTAGCACAGCTTCCAGGCGCTGGTG	1202
Qy	1104	GTACAGGCCAGTCCCGAAGTTTCACTATAGAGGGCGGAACTGTGCGGCCGCTACAT	1163
Db	1203	GTACAGGCCAGTCCCGAAGTTTCACTATAGAGGGCGGAACTGTGCGGCCGCTACAT	1262

QY	1164	TCACATTCGGAGAGGGGGATTTTACTACCAACCGCGCCATTTTTCGGAGAACGCGCTACTCA	1223
Db	1263	TCACATTCGGAGAGGGGGATTTTACTACCAACCGCGCCATTTTTCGGAGAACGCGCTACTCA	1322
QY	1224	CATTCTCAACCTCACACGGAGTCTTTTAAGGGGATCCTCCAGGACAAGAACATGACAT	1283
Db	1323	CATTCTCAACCTCACACGGAGTCTTTTAAGGGGATCCTCCAGGACAAGAACATGACAT	1382
QY	1284	AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCCCTTTTCAACAACCCAC	1343
Db	1383	AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCCCTTTTCAACAACCCAC	1442
QY	1344	TAAATCTCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCGTCGCTTCAGATATTAA	1403
Db	1443	TAAATCTCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCGTCGCTTCAGATATTAA	1502
QY	1404	RAGTCTCAAGGTACCTTGGCAGACAAAAGATATAATTTGGTTAGAANAATATGCTCTGACT	1463
Db	1503	RAGTCTCAAGGTACCTTGGCAGACAAAAGATATAATTTGGTTAGAANAATATGCTCTGACT	1562
QY	1464	TCAAAATCTGTATGGAACCTTGACACTATTTCCT	1495
Db	1563	TCAAAATCTGTATGGAACCTTGACACTATTTCCT	1594
RESULT	8		
LOCUS	AR066335	3450 bp	DNA linear PAT 29-SEP-1999
DEFINITION	Sequence 9 from patent US 5849991.		
ACCESSION	AR066335		
VERSION	AR066335.1	GI:5996551	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 3450) d'Aplice,A.J.F., Pearse,M.J., Robins,A.J., Crawford,R.J. and Rathjen,P.D.		
TITLE	Mice homozygous for an inactivated .alpha. 1,3-galactosyl transferase gene		
JOURNAL	Patent: US 5849991-A 9 15-DEC-1998;		
FEATURES	Location/Qualifiers 1..3450 /organism="unknown"		
BASE COUNT	954 a 762 c 826 g 908 t		
ORIGIN			
Query Match	93.9%	Score 1408.4;	DB 6; Length 3450;
Best Local Similarity	96.8%;	pred. No. 0;	
Matches 1483;	Conservative 0;	Mismatches 11;	Indels 38; Gaps 3;
QY	1	CCTTCCCTTTGAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT	60
Db	64	CCTTCCCTTTGAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT	123
QY	61	CTCAGGCTCTGGGAGTTGGAACCTT-GTACCTTCTTTCCTCTGCTGAGCCCTCGCTTCT	119
Db	124	CTCAGGCTCTGGGAGTTGGAACCTTGTACCTTCTTTCCTCTGCTGAGCCCTGCTTCT	183
QY	120	TAGGCAGGCCAGAGCTCGACAGAACTCGGTTGCTTTTGTCTGTTTGGAGGGAACACA	179
Db	184	TCGGCAGGCCAGAGCTGACAGAAGCTCGGTGCTTTTGTCTGTTTGGAGGGAACACA	243
QY	180	GCTGACGATAGGCTGACTTTTGAACCTCAAGAGATCTGCTTACCCAGTCTCTCGGAATTA	239
Db	244	GCTGACGATAGGCTGACTTTTGAACCTCAAGAGATCTGCTTACCCAGTCTCTCGGAATTA	303
QY	240	RAGGCTCTACTACATTTGCGCTGGACCTTAGATTTTTCATGATCACTATGCTTCAAGATCT	299
Db	304	RAGGCTCTACTAC -CTTGCTGGACCTTAGATTTTTCATGATCACTATGCTTCAAGATCT	362
QY	300	CCATGTCACCAAGATCTCCATGTCAAGATCAAGTCAAGAACCAAGTCTTGATCTCCTCAAG	359



|||||  
Db 363 CCATGTCACAGAGTCTCCATGTCACAGATCCAAAGTCAGAAACAGCTCTCCATCTCTCAAG 422  
|||||  
QY 360 ATCTGGATCACAGAGAGAAATAATGAATGTCACAGGGAAGATGAATCTGTTGATGCTGAT 419  
|||||  
Db 423 ATCTGGATCACAGAGAGAAATAATGAATGTCACAGGGAAGATGAATCTGTTGATGCTGAT 482  
|||||  
QY 420 TGTCTCAACCGTGGTGTCTGCTGTTTGGGAAATATGTCACAG----- 461  
|||||  
Db 483 TGTCTCAACCGTGGTGTCTGCTGTTTGGGAAATATGTCACAGCCAGCGCTCTTCTT 542  
|||||  
QY 462 -----AATTCACAGAGTGTGTGAGACAGATGCGCAGAGGACTGGTG 503  
|||||  
Db 543 GTGGATATATCACACAAAATTCACAGAGTGTGTGAGACAGATGCGCAGAGGACTGGTG 602  
|||||  
QY 504 GTTCCCAAGCTGCTTTAAAAATGGGACCCACAGTTATCAAGAGACAACTGACAGGACG 563  
|||||  
Db 603 GTTCCCAAGCTGCTTTAAAAATGGGACCCACAGTTATCAAGAGACAACTGACAGGACG 662  
|||||  
QY 564 GAGAGAAAAGGGTAGAAATGGAGATCGCATTTGAAGAGCCCTACGCTATGGGACTGGTTCAA 623  
|||||  
Db 663 GAGAGAAAAGGGTAGAAATGGAGATCGCATTTGAAGAGCCCTACGCTATGGGACTGGTTCAA 722  
|||||  
QY 624 TCCAAAGAACCGCCCGGATGTTTTGACAGTGACCTGACCCGCGTGAAGCGCGGATTTGTGGGA 683  
|||||  
Db 723 TCCAAAGAACCGCCCGGATGTTTTGACAGTGACCTGACCCGCGTGAAGCGCGGATTTGTGGGA 782  
|||||  
QY 684 AGGCACCTTATGACACAGCTCTGCTGGGAAAGTACTACGCCACACAGAACTCACTGTGGG 743  
|||||  
Db 783 AGGCACCTTATGACACAGCTCTGCTGGGAAAGTACTACGCCACACAGAACTCACTGTGGG 842  
|||||  
QY 744 GCTGACAGTGTTCCTGTGGGAAAGTACATTTAGACATTTACTTGAAGACATTTCTGGAGTC 803  
|||||  
Db 843 GCTGACAGTGTTCCTGTGGGAAAGTACATTTAGACATTTACTTGAAGACATTTCTGGAGTC 902  
|||||  
QY 804 TGCTGACATGACTCATGTTGGCCATCGGGTCATATTTTAGCTCATGATACACACAC 863  
|||||  
Db 903 TGCTGACATGACTCATGTTGGCCATCGGGTCATATTTTAGCTCATGATAGATGACAC 962  
|||||  
QY 864 CTCGGGATGCTGCTGACCTGAAACCTCTACATTTCTTCAAGTCTTTGAGATCAG 923  
|||||  
Db 963 CTCGGGATGCTGCTGACCTGAAACCTCTACATTTCTTCAAGTCTTTGAGATCAG 1022  
|||||  
QY 924 GTCTGAGAGAGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGACACAT 983  
|||||  
Db 1023 GTCTGAGAGAGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGACACAT 1082  
|||||  
QY 984 CCTGGCCACATCCAGCACAGTTCGACTTCCTCTCTGCTGATGACGCTGGATCAAGTCTT 1043  
|||||  
Db 1083 CCTGGCCACATCCAGCACAGTTCGACTTCCTCTCTGCTGATGGAGTGGATCAAGTCTT 1142  
|||||  
QY 1044 TCAAGACAACCTTCGGGGTGAAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCTGGTG 1103  
|||||  
Db 1143 TCAAGACAACCTTCGGGGTGAAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCTGGTG 1202  
|||||  
QY 1104 GTACAAGGCCAGTCCCGAGAGTTCACTATGAGAGCGGGAACTGTGCGCCCGCTACAT 1163  
|||||  
Db 1203 GTACAAGGCCAGTCCCGAGAGTTCACTATGAGAGCGGGAACTGTGCGCCCGCTACAT 1262  
|||||  
QY 1164 TCCATTCCGAGAGGGGATTTTACTACACAGCGGGCCATTTTTCGAGGAACGCTACTCA 1223  
|||||  
Db 1263 TCCATTCCGAGAGGGGATTTTACTACACAGCGGGCCATTTTTCGAGGAACGCTACTCA 1322  
|||||  
QY 1224 CATTTCTCAACCTTCACAGGAGTGTCTTAAAGGGATCTCTCAGGACAAAGAAACATGACAT 1283  
|||||  
Db 1323 CATTTCTCAACCTTCACAGGAGTGTCTTAAAGGGATCTCTCAGGACAAAGAAACATGACAT 1382  
|||||  
QY 1284 AGAAGCCCACTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCAC 1343  
|||||  
Db 1383 AGAAGCCCACTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCAC 1442  
|||||  
QY 1344 TAAATCTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCGTGCCTTCAGATATTAA 1403  
|||||

Db 1443 TAAATCTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCGCTCCCTTCAGATATTAA 1502  
QY 1404 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATTAATTTGTTGTAGAAATTAATCTCTGACT 1463  
|||||  
Db 1503 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATTAATTTGTTGTAGAAATTAATCTCTGACT 1562  
|||||  
QY 1464 TCAATTTGTGATGGAACCTTTGACACTATTCT 1495  
|||||  
Db 1563 TCAATTTGTGATGGAACCTTTGACACTATTACT 1594  
|||||  
RESULT 9  
BC006810 3315 bp mRNA linear ROD 12-JUL-2001  
LOCUS Mus musculus, clone MGC:11545 IMAGE:3154188, mRNA, complete cds.  
DEFINITION BC006810  
ACCESSION BC006810.1 GI:13905053  
VERSION MGC.  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 3315)  
Strausberg,R.  
Direct Submission  
Submitted (27-APR-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: http://mgc.ncl.nih.gov  
Contact: MGC help desk  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
Contact: villalona@bcm.tmc.edu.  
Villalona, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,  
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
Muzny, D.M., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAK Plate: 16 Row: C Column: 15  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 193419.  
FEATURES  
Location/Qualifiers  
1..3315  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="MGC:11545 IMAGE:3154188"  
/tissue\_type="Mammary tumor. Brcal-/-fl; MMTV-Cre model. 10  
months old, gross tissue."  
/clone\_lib="NCI\_CGAP\_Mam3"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
252..1370  
/codon\_start=1  
/product="Unknown (protein for MGC:11545)"  
/protein\_id="AAH06810.1"  
/db\_xref="GI:13905054"  
/translation="MITMLQDLHVKNKISMRSKSETSLPSSRSQSOFKIMNVKGVIL  
LMLIVSTVVVFWEYVNRTHSTQEDNVEGRNRKNGRDIEEPQLDWPNKRPDVL  
TYPWKPAPIVWEGTYDTALLEKYATOKLTGLTVFAVGKYIEHYLEDLESADNPFM  
VGRHVFYVYIDDTSRMPVHLNPLSLQVFETIRSEKRMQDI SHMRKTI GEHILAH  
QHEVDPLFCMDVDQVDFQDFVETLGLVAQLQAWKYKASPEKFTYERRLSAAYIPF  
GEGEDVYHAAITGGTPTHTLNLTRFCKGLQDKKHIDIEAQWHDHSLNKYFLFNKPT  
KILSPCYCHDYQIGLPSDIKSYKVAQTKREYNLVRNV"  
937 a 719 c 782 g 877 t  
BASE COUNT  
ORIGIN

Query Match	88.4%	Score 1326.4	DB 10	Length 3315	
Best Local Similarity	95.4%	Pred. No. 0			
Matches 1403	Conservative	0	Mismatches 1	Indels 66	Gaps 1
Qy	26	ATGAGAAGTACCGGATCTTGCTGAAGACTCGCGCTCTCAGGCTCTGGGAGTTGGAAACCT	85		
Db	1	ATGAGAAGTACCGGATCTTGCTGAAGACTCGCGCTCTCAGGCTCTGGGAGTTGGAAACCT	60		
Qy	86	GTACCTTTCCTTTCTCTGCTGAGCCCTGCCCTTTCCTTTAGCAGCGCAGAGCTCGACAGAACT	145		
Db	61	GTACCTTTCCTTTCTCTGCTGAGCCCTGCCCTTTCCTTTAGCAGCGCAGAGCTCGACAGAACT	120		
Qy	146	CGGTTGCTTTCCTCTTTCCTTTGGAGGGAACACAGCTGACGATGAGGCTGACTTTGAACT	205		
Db	121	CGGTTGCTTTCCTCTTTCCTTTGGAGGGAACACAGCTGACGATGAGGCTGACTTTGAACT	180		
Qy	206	CAAGAGATCTGCTTTACCCCACTCTCCTGGAAATTAAGGCCCTTACTACATTTGCTGGAC	265		
Db	181	CAAGAGATCTGCTTTACCCCACTCTCCTGGAAATTAAGGCCCTTACTACATTTGCTGGAC	240		
Qy	266	CTAAGATTTTCATGATCACTATGCTTCAGATCTCCATGTCCATGTCAACAAGATCTCCATGTCAA	325		
Db	241	CTAAGATTTTCATGATCACTATGCTTCAGATCTCCATGTCCATGTCAACAAGATCTCCATGTCAA	300		
Qy	326	GATCCAAGTCAGAACAAGTCTTCATCCTCAAGATCTGGATCAAGGAGAAAATAATGA	385		
Db	301	GATCCAAGTCAGAACAAGTCTTCATCCTCAAGATCTGGATCAAGGAGAAAATAATGA	360		
Qy	386	ATGTCAAGGGAAAAGTAATCCTGTTGATGCTGATGCTCAACCGTGTTGTCTGCTTTT	445		
Db	361	ATGTCAAGGGAAAAGTAATCCTGTTGATGCTGATGCTCAACCGTGTTGTCTGCTTTT	420		
Qy	446	GGGAATATGTCAACAGAAATCCAGAGTTGCTGAGAACAGATGGCAGAGGACTGGTGGT	505		
Db	421	GGGAATATGTCAACA-----	435		
Qy	506	TCCCAAGCTGGTTTAAAAATGGGACCCACAGTTATCAAGAAGACAACTAGAAGACGGA	565		
Db	436	-----GGACCCACAGTTATCAAGAAGACAACTAGAAGACGGA	474		
Qy	566	GAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAATC	625		
Db	475	GAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAATC	534		
Qy	626	CAAGAACCGCCCGATCTTTTGACGTGACCCCGTGAAGCGCCGATTTGTGGGAAG	685		
Db	535	CAAGAACCGCCCGATCTTTTGACGTGACCCCGTGAAGCGCCGATTTGTGGGAAG	594		
Qy	686	GCACCTATGACACAGCTCTGCTGGAAAAAGTACTACGCCACACAGAACTCAGCTGTGGGC	745		
Db	595	GCACCTATGACACAGCTCTGCTGGAAAAAGTACTACGCCACACAGAACTCAGCTGTGGGC	654		
Qy	746	TGACAGTCTTTGCTTGGGAAGTACATTTAGCATTACTTAGAAGCTTTCTGGAGTCTG	805		
Db	655	TGACAGTCTTTGCTTGGGAAGTACATTTAGCATTACTTAGAAGCTTTCTGGAGTCTG	714		
Qy	806	CTGACATGTACTTCATGGTTGGCCATCGGGTCATATTTTAGCTCATGATAGACGACACT	865		
Db	715	CTGACATGTACTTCATGGTTGGCCATCGGGTCATATTTTAGCTCATGATAGACGACACT	774		
Qy	866	CCCGGATGCCCTGCTGTGACACCTGAACCCCTCTACATTTACAAGCTTTTGAGATCAGGT	925		
Db	775	CCCGGATGCCCTGCTGTGACACCTGAACCCCTCTACATTTACAAGCTTTTGAGATCAGGT	834		
Qy	926	CTGAGAGAGGTGGCAGGATATCAGCATGATCGCATCAAGACCAATTCGGGAGCACATCC	985		
Db	835	CTGAGAGAGGTGGCAGGATATCAGCATGATCGCATCAAGACCAATTCGGGAGCACATCC	894		
Qy	986	TGGCCACATCCAGCACGAGTCTGACTTCTCTTCTGATGACGCTGGATCAAGTCTTTC	1045		
Db	895	TGGCCACATCCAGCACGAGTCTGACTTCTCTTCTGATGACGCTGGATCAAGTCTTTC	954		

Qy	1311	CCTCAACAAATACATTCTCTTTTCAACAAACCCACTAAAAATCCTATCTCCAGAGTATTGCTG	1370			
Db	979	TCTAAACAAGTATTCTCTTCTCAACAAACCCACTAAAAATCTTATCCCAAGAAATCTGCTG	1030			
Qy	1371	GGACTATCAGATAGGCTGCGTTTCAGATATTAAGTGTCAAGGTAGCTTGGCAGACAAA	1430			
Db	1039	GGATTATCATATAGGCATGCTCTGGGATATATAGGATTTGTCAAGATAGCTTGGCAGAAAA	1090			
Qy	1431	AGAGTATAAATTTGGTTAGAANAATGTCTGACCTTCAAATTTGTGATGGAAACTT	1483			
Db	1099	AGAGTATAAATTTGGTTAGAANAATCATCTGACTTTAAATTTGTGCCAGCAGTTT	1151			
RESULT 11						
AR048516						
LOCUS	AR048516	1423 bp	DNA			
DEFINITION	Sequence 2 from patent US 5821117.					
ACCESSION	AR048516					
VERSION	AR048516.1	GI:5970859				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1423)					
AUTHORS	Sandrin,M.S. and McKenzie,I.F.C.					
TITLE	Xenotransplantation therapies					
JOURNAL	Patent: US 5821117-A 2 13-OCT-1998;					
FEATURES	Location/Qualifiers					
source	1..1423					
BASE COUNT	413 a	305 c	348 g			
ORIGIN	/organism="unknown"					
Query Match 48.4%; Score 726.6; DB 6; Length 1423;						
Best Local Similarity 78.8%; Pred. No.3e-184;						
Matches 893; Conservative 0; Mismatches 234; Indels 6; Gaps						
Qy	354	CTCAAGATCTGGATCAGGAGAAATATGAATGTCAAGGGAAAGTAATCCTGTGTGAT	413			
Db	22	CCAGCTTCGCGCTCAGGAGAAATATGAATGTCAAGGAAAGTGGTCTGTCTCAAT	81			
Qy	414	GCTGATTGTCTCAACCGTGGTTGCTGTTTTCGGGAATATGTCAACAGAAATTCACAGGT	473			
Db	82	GCTGCTTGTCTCAACTGTAATGCTGTGTTTTCGGGAATACATCAACAGAAACCCAGAGCT	141			
Qy	474	TGGTCAACACAGATGGCAGAAAGACGTGGTGGTTCCTCCACAGCTGGTTTAAAAATGGGACCA	533			
Db	142	TGG---CAGCAGTGTCTCAGAGGGGCTGGTGGTTTTCGAGCTGGTTTAAACAATGGGACTCA	198			
Qy	534	CAGTTATTCAAGACACAACCTAGAAAGCAGCAGAGAAAGG---GTAGAAATGGAGATCG	590			
Db	199	CAGTTACCAAGACAGACAGACCTATAGCACAAGAAAGGAACAAGAAAGAGACAA	258			
Qy	591	CAITTAAGAGCCTCAGCTATGGACGTGGTTCAATCCAAAGAACCGCCGGATGTTTTTCAC	650			
Db	259	CAGAGGAGAGCTCCGCTAGTGGACTGGTTTAACTCTGAGAAAGCCAGAGGTGCTGAC	318			
Qy	651	AGTCACCCGTGGAAAGCCCGGATTTGTGGGGAAGGACCTTATGACACAGACTCTGCTGGA	710			
Db	319	CATTAACACAGATGGAAGGCTCCAGTGGTGTGGGAAGGACCTTACAACAGACAGCGTCTAGA	378			
Qy	711	AACTACTACGCCACAGAAACCTACTGTGGGGCTGTACAGCTGTTTGCCTGTCGGGAAGTA	770			
Db	379	TAATTATTATGCCAACAGAAATATACCGTGGGCTGTGACGGTTTTTGCCTGTCCGAAGATA	438			
Qy	771	CATTGAGCATTTACTTAGAAGACTTTCTGGAGCTCTGCTGACATGTACTTTCATGTTGGCCA	830			
Db	439	CATTGAGCATTTACTTGAGGAGTTCTTAATATCTGCAATACATACTTTCATGTTGGCCA	498			
Qy	831	TCGGGTCAATTTTACGTCATATAGACACACCTCCCGGATGCCTGTCTGTCGACCTGAA	890			
Db	499	CAAACTCATCTTTTACATCATGGTGGATGATATCTCAGGATGCCCTTTCATAGAGCTGGG	558			



QY 1306 AGCCACCTCAACAAATACCTTCCCTTTTCAACAAACCACATAAATCCTATCTCCAGAGTAT 1365  
|||||  
Db 935 AGCCACCTCAACAAATACCTTCCCTTTTCAACAAACCACATAAATCCTATCTCCAGAGTAT 994  
QY 1366 TCGTGGGACTATCAGATAGGCGCTTCCAGATATTTAAAGTGTCAAGGTAGCTTGGCAG 1425  
|||||  
Db 995 TCGTGGGACTATCAGATAGGCGCTTCCAGATATTTAAAGTGTCAAGGTAGCTTGGCAG 1054  
QY 1426 ACAAAGAGTATATTTGGTTAGAAATAATGCTGACTTCAAAATGCTGATGGAACCTTGA 1485  
|||||  
Db 1055 ACAAAGAGTATATTTGGTTAGAAATAATGCTGACTTCAAAATGCTGATGGAACCTTGA 1114  
QY 1486 CACTATTCTT 1495  
|||||  
Db 1115 CACTATTACT 1124  
RESULT 13  
AX128258  
LOCUS AX128258 3240 bp DNA linear PAT 15-MAY-2001  
DEFINITION Sequence 25 from Patent WO0130992.  
ACCESSION AX128258  
VERSION AX128258.1 GI:14134779  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 3240)  
AUTHORS Koike,C.  
TITLE \_g(a)1-3 galactosyltransferase gene and promoter  
JOURNAL Patent: WO 0130992-A 25 03-MAY-2001;  
UNIV. PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION  
(US)  
FEATURES  
source Location/Qualifiers  
1..3240  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
1..369  
misc\_feature 370..3010  
/note="exon 9"  
terminator 1088..1090  
3'UTR 1091..3010  
intron 3011..3240  
BASE COUNT 901 a 694 c 727 g 878 t 40 others  
ORIGIN  
Query Match 48.3%; Score 725.2; DB 6; Length 3240;  
Best Local Similarity 99.6%; Pred. No. 8.1e-184;  
Matches 727; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 766 AAGTACATTGAGGATTACTTAGAAGACTTCTGGAGCTTCTGACATGACTTTCATGGTT 825  
|||||  
Db 395 AGGTACATTGAGGATTACTTAGAAGACTTCTGGAGCTTCTGACATGACTTTCATGGTT 454  
QY 826 GGCCATCGGTCATATTTACGTCATGATAGACGACACCTCCCGGATGCTGTCGTGCAC 885  
|||||  
Db 455 GGCCATCGGTCATATTTACGTCATGATAGATGACACCTCCCGGATGCTGTCGTGCAC 514  
QY 886 CTGAACCCCTACATCTCTTACAAGCTTTTGAGATCAGGTCTGAGAAGAGGTGGCAGGAT 945  
|||||  
Db 515 CTGAACCCCTACATCTCTTACAAGCTTTTGAGATCAGGTCTGAGAAGAGGTGGCAGGAT 574  
QY 946 ATCAGCATGATGGCATGAAGACCATTTGGGGAGCAGCATCTCGGCCACATCCAGCACGAG 1005  
|||||  
Db 575 ATCAGCATGATGGCATGAAGACCATTTGGGGAGCAGCATCTCGGCCACATCCAGCACGAG 634  
QY 1006 GTCGACTTCTCTTCGATGACGCTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAA 1065  
|||||  
Db 635 GTCGACTTCTCTTCGATGACGCTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAA 694

QY 1066 ACTCTGGGCCAGCTGGTAGCAGACTCCAGGCTGGTGGTACAAAGCCAGTCCCGAGAAG 1125  
|||||  
Db 695 ACTCTGGGCCAGCTGGTAGCAGACTCCAGGCTGGTGGTACAAAGCCAGTCCCGAGAAG 754  
QY 1126 TTACACCTATGAGAGGGGGAACTGTGCGCGCGGTACATTTCCATTTCCGAGAGGGGGATTTT 1185  
|||||  
Db 755 TTACACCTATGAGAGGGGGAACTGTGCGCGCGGTACATTTCCATTTCCGAGAGGGGGATTTT 814  
QY 1186 TACTACACCGCGCCATTTTGGAGGAAGCCCTACTTCACATTTCTCAACCTTCACACGGGAG 1245  
|||||  
Db 815 TACTACACCGCGCCATTTTGGAGGAAGCCCTACTTCACATTTCTCAACCTTCACACGGGAG 874  
QY 1246 TCGTTTAAAGGGATCCTCCAGGACAAAGAACATGACATAGAACCCAGTGGCATGATGAG 1305  
|||||  
Db 875 TCGTTTAAAGGGATCCTCCAGGACAAAGAACATGACATAGAACCCAGTGGCATGATGAG 934  
QY 1306 AGCCACCTCAACAAATACCTTCCCTTTTCAACAAACCACATAAATCCTATCTCCAGAGTAT 1365  
|||||  
Db 935 AGCCACCTCAACAAATACCTTCCCTTTTCAACAAACCACATAAATCCTATCTCCAGAGTAT 994  
QY 1366 TCGTGGGACTATCAGATAGGCGCTTCCAGATATTTAAAGTGTCAAGGTAGCTTGGCAG 1425  
|||||  
Db 995 TCGTGGGACTATCAGATAGGCGCTTCCAGATATTTAAAGTGTCAAGGTAGCTTGGCAG 1054  
QY 1426 ACAAAGAGTATATTTGGTTAGAAATAATGCTGACTTCAAAATGCTGATGGAACCTTGA 1485  
|||||  
Db 1055 ACAAAGAGTATATTTGGTTAGAAATAATGCTGACTTCAAAATGCTGATGGAACCTTGA 1114  
QY 1486 CACTATTCTT 1495  
|||||  
Db 1115 CACTATTACT 1124  
RESULT 14  
A69346  
LOCUS A69346 1092 bp DNA linear PAT 06-MAY-1999  
DEFINITION Sequence 3 from Patent WO9802453.  
ACCESSION A69346  
VERSION A69346.1 GI:4760173  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 1092)  
AUTHORS Gourlet,P., Robberecht,P., Vandermeers,A. and Woelbroeck,M.  
TITLE PEPTIDIC LIGANDS HAVING A HIGHER SELECTIVITY FOR THE VIP1 RECEPTOR  
JOURNAL THAN FOR THE VIP2 RECEPTOR  
Patent: WO 9802453-A 3 22-JAN-1998;  
UNIV BRUXELLES (BE)  
FEATURES  
source Location/Qualifiers  
1..1092  
/organism="unidentified"  
/db\_xref="taxon:32644"  
<1..1092  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB42481.1"  
/db\_xref="GI:4760174"  
/translation="NSKIMNVKGRVLSMLLYSTVMVVFWEYINRNPEVGSQAQRGHW  
PFSWFNNGTHSYHEEDAIGNEKEQRKEDNKGELPLVDNFNPKRPEVVTITRMKAPV  
WVEGTYNRAVLNDNYAKQITVGLTVLAVGRY IEHYLEEFLLISANTYVMVGHVPIFY  
MVDDISRMPLIELGRSPKVPKVEIKSEKRWDISMMRMKTIGEHILAHIQHVDPLFC  
MDVDQVFNQFVETLGOSVAOLOAWYKAHPDEPTYERRKESAAY IPFGQGDFFYYHA  
ALFGCTPTQVNLNITQECFKGILQDKENDIEAEWHDESHLNKFLNKPTKILSPYCVH  
DYHIGMSVDIRIVKIAWQKKEYNLRNNI"  
BASE COUNT 325 a 228 c 279 g 260 t  
ORIGIN  
Query Match 47.1%; Score 706; DB 6; Length 1092;  
Best Local Similarity 79.3%; Pred. No. 9.9e-179;  
Matches 864; Conservative 0; Mismatches 220; Indels 6; Gaps 2;

```
QY 375 GAAATAATGAATGTCAGGGAAGTAATCCCTGTTGATGCTGANTGCTCTCAACCGTGGT 434
Db 6 GAAATAATGAATGTCAGGGAAGTGGTCTGTCAATGCTGCTCTCAACTGTAAT 65
QY 435 TGTGCTGTTTGGGAATATGTCACAGAAATCCAGAGAGTTGGTGAGAACAGATGGCAGAA 494
Db 66 GGTGTGCTGTTTGGGAATATGTCACAGAAATCCAGAGAGTTGG---CAGCAGTGTCTCAGAG 122
QY 495 GGACTGTGTTTCCCAAGCTGGTTTAAATAATGGGACCCACAGTTATCAAGAACAACACGT 554
Db 123 GGGCTGTGTTTCCGAGCTGGTTTACAAATGGGACTCACAGTTACCAACAGAGAGA 182
QY 555 AGAGGACCGAGAGAAAAGG---GTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATG 611
Db 183 CGCTATAGGCAAGCAAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAGAGCTCCACTAGT 242
QY 612 GGACTGTGTTCAATCCAAAGACCGCCGGATGTTTTCACAGTGCACAGTGCAGAGCGCC 671
Db 243 GGGCTGTGTTTAACTCGAGAAAGCCAGAGGCTGCGTACCATACCCAGATGGAGAGGCTCC 302
QY 672 GATTGTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAAGTACTACGCCACACAGAA 731
Db 303 AGTGGTATGGAAGGCACTTACAACAGAGCGCTCTTAGATAATATTATGCCAAACAGAA 362
QY 732 ACTCACTGTGGGCTGACAGTGTGTTGCTGTGGGAAAGTACATTTAGAGCACTTACTTAGA 791
Db 363 AATTACCGTGGGCTGACAGTGTGTTGCTGTGGGAAAGTACATTTAGAGCACTTACTTTGAG 422
QY 792 CTTCCTGAGCTGCTGACATGACTTTCATGTTGGCCATCGGCTCATATTTTACGTCAT 851
Db 423 GTTCTTAATATCTGCAAAATACATCTTCATGTTGGCCAAAGTCACTCTTTTACATCAT 482
QY 852 GATAGAGGACACCTCCCGGATGCTGCTGCACCTGAAACCCCTACATTTCTTTTACAAGT 911
Db 483 GGTGGATGATATCTCCAGGATGCTTTGATAGAGCTGGGCTCTGCGCTCTTTAAAGT 542
QY 912 CTTTGATGATCAGTCTGAGAAAGTGGCAGGATATCAGCATGATGCGCATGAAGACCAT 971
Db 543 GTTTGATGATCAAGTCCGAGAAAGTGGCAAGACATCAGCATGATGCGCATGAAGACCAT 602
QY 972 TGGGAGCACATCTGGCCACATCCAGCACGAGGTGCGACTTCTCTTCGATGGAGCT 1031
Db 603 CGGGAGCACATCTGGCCACATCCAGCACGAGGTGCGACTTCTCTTCGATGGAGCT 662
QY 1032 GGATCAAGTCTTTCAAGACAACTTCGGGGTGGAAACTCTGGGCGAGCTGTGAGCACAGCT 1091
Db 663 GGATCAGGCTCTCCAAACAACTTTGGGGTGGAGACCTTGGCCAGCTGGGCTGCTCAGCT 722
QY 1092 CCAGGCTGTGTGATCAAGGCCAGTCCGAGAAAGTTTACCTATGAGAGCGGGAAGTCT 1151
Db 723 ACAGGCTGTGTGATCAAGGCCATCTGACGAGTTTACCTACGAGAGCGGGAAGGAGTC 782
QY 1152 GGGCGGTGATCTTCCATTCGAGAGGGGGATTTTACTACACGCGGCGCATTTTGGAGG 1211
Db 783 CGCAGCTACATTTCCGTTTGGCCAGGGGGATTTTATACACGCGGCGCATTTTGGGG 842
QY 1212 AACCCCTACTACATTTCTCAACCTCACAGGAGTGGTTTAAAGGGATCTCTCCAGGACAA 1271
Db 843 AACCCCACTCAGTTCTTAAACATCACTCAGGAGTGGTTTCAAGGGAATCTCTCCAGGACAA 902
QY 1272 GAACATGACATAGAGCCCAAGTGGATGATGAGAGCCACTCAACAAATACTTCTCTTT 1331
Db 903 GGAATATGATATAGAGCCGAGTGGCATGATGAAGCCATCTAAACAAATATTTCTCTCT 962
QY 1332 CAACAAACCCACTAAATCTCTATCCAGAGTATGCTGGGACTATCAGATAGCCCTGCC 1391
Db 963 CACAAACCCCACTAAATCTTATCCCAAGAAATATCTGCTGGGATATCATATAGGCAATGTC 1022
QY 1392 TTCAGATATTTAAAGTGTCAAGGTAGTGTGGCAGACAAAAGATATATTTTGGTTAGAAA 1451
Db 1023 TGTGGATATTTAGGATTTGTCAGATAGCTTGGCAGAAAAGAGATATATTTTGGTTAGAAA 1082
QY 1452 TAATGCTGTA 1461
```

Db 1083 TAACATCTGA 1092

RESULT 15

A72023 LOCUS A72023 1092 bp DNA linear PAT 11-MAY-1999  
DEFINITION Sequence 3 from Patent WO9803653.  
ACCESSION A72023  
VERSION A72023.1 GI:4808049  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 1092)  
AUTHORS Pourcel,C., Soullilou,J. and Vanhove,B.  
TITLE METHOD FOR PREPARING TRANSGENIC NON-HUMAN MAMMALIAN ORGANS FOR  
TRANSPLANTATION TO HUMANS, AND NUCLEOTIDE SEQUENCES THEREFOR  
JOURNAL Patent: WO 9803653-A 29-JAN-1998;  
COMMENT INST NAT SANTE RECH MED (FR)  
FEATURES Other publication FR 2751346 19980123.  
Location/Qualifiers  
1..1092  
/organism="unidentified"  
/db\_xref="taxon:32844"  
<1..1092  
/notes="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB42704.1"  
/db\_xref="GI:4808050"

CDS

325 a 228 c 279 g 260 t  
BASE COUNT  
ORIGIN  
Query Match 47.1%; Score 706; DB 6; Length 1092;  
Best Local Similarity 79.3%; Pred. No. 9.9e-179;  
Matches 864; Conservative 0; Mismatches 220; Indels 6; Gaps 2;  
QY 375 GAAATAATGAATGTCAGGGAAGTAATCCCTGTTGATGCTGANTGCTCTCAACCGTGGT 434  
Db 6 GAAATAATGAATGTCAGGGAAGTGGTCTGTCAATGCTGCTCTCAACTGTAAT 65  
QY 435 TGTGCTGTTTGGGAATATGTCACAGAAATCCAGAGGTTGGTGAGAACAGATGGCAGAA 494  
Db 66 GGTGTGCTGTTTGGGAATATGTCACAGAAATCCAGAGAGTTGG---CAGCAGTGTCTCAGAG 122  
QY 495 GGACTGTGTTTCCCAAGCTGGTTTAAATAATGGGACCCACAGTTATCAAGAACAACACGT 554  
Db 123 GGGCTGTGTTTCCGAGCTGGTTTACAAATGGGACTCACAGTTACCAACAGAGAGA 182  
QY 555 AGAGGACCGAGAGAAAAGG---GTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATG 611  
Db 183 CGCTATAGGCAAGCAAGGAAGAAAGAAAGAAAGAAAGAGAGAGCTCCACTAGT 242  
QY 612 GGACTGTGTTCAATCCAAAGACCGCCGGATGTTTTCACAGTGCACAGTGCAGAGCGCC 671  
Db 243 GGGCTGTGTTTAACTCGAGAAAGCCAGAGGCTGCGTACCATACCCAGATGGAGAGGCTCC 302  
QY 672 GATTGTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAAGTACTACGCCACACAGAA 731  
Db 303 AGTGGTATGGAAGGCACTTACAACAGAGCGCTCTTAGATAATATTATGCCAAACAGAA 362  
QY 732 ACTCACTGTGGGCTGACAGTGTGTTGCTGTGGGAAAGTACATTTAGAGCACTTACTTAGA 791  
Db 363 AATTACCGTGGGCTGACAGTGTGTTGCTGTGGGAAAGTACATTTAGAGCACTTACTTTGAG 422

Qy	792	CTTTCTGGAGTCTGCTGACATGTACTTATGTTTGCCCATCGGGTCATATTTTTACGTCAT	851
Db	423	GTTCTTAATATCTGCAAAATACATACTTCATGTTTGCCCACAAAGTCATCTTTTACATCAT	482
Qy	852	GATGACGCACACCTCCGGATGCTGCTGTCGACCCTGAACCTCTACATCTCCTTACAAGT	911
Db	483	GGTGATGATATCTCCAGGATGCCTTTGATAGAGCTGGGTCTCTGCGCTCCTTTAAGT	542
Qy	912	CTTTTGAGATCAGGCTTGAGAAGAGGTGGCAGGATATCAGCATGATGGCATGAAGACCAT	971
Db	543	GTTTGAGATCAAGTCCGAGAAGAGGTGGCAAGACATCAGCATGATGGCATGAAGACCAT	602
Qy	972	TGGSGAGCACATCTCGCCACATCCAGCACGAGTGCACATTCTCTTCGTCATGGACGT	1031
Db	603	CGGGGAGCACATCTCGCCACATCCAGCACGAGGTGCATCTCTCTTCGATGGACGT	662
Qy	1032	GGATCAAGTCTTTCAAGACAACCTTCGGGTGGAAACTCTGGGCCAGCTGGTAGCACAGCT	1091
Db	663	GGATCAGGCTTCCAAACAACCTTTGGGTGGAGACCTGGGCCAGTCGGTGGCTCAGCT	722
Qy	1092	CCAGGCTGTGGTACAAAGCCNAGTCCCAGAAAGTTCACTTATGAGAGGCGGAACCTGTC	1151
Db	723	ACAGGCTGTGGTACAAAGCACATCTTCACGAGTTCACTTACGAGAGGCGGAAGGAGTC	782
Qy	1152	GGCCGCGTACATTCCTATTCGGAGAGGGGATTTTTACTACCACCGCGGCATTTTTTGAGG	1211
Db	783	CGCAGCCTACATTTCCGTTTGGCCAGGGGATTTTTATTACCACGACGCCATTTTTGGGG	842
Qy	1212	AACGCCCTACTACATTTCTCAACGTCACAGGAGGTGCTTTAAGGGGATCTCTCAGGACAA	1271
Db	843	AACACCCACTCAGGTTTCTAAATCACTACGAGGTGCTTCAAGGGAA'TCCTCCAGSACAA	902
Qy	1272	GAACAATGACATAGAAGCCAGTGGCATGATCAGAGGCACCTCAACAATAACTTCTCTTTT	1331
Db	903	GGAAATGACATAGAAGCCAGTGGCATGATCAAGGCCATCTAAACAGTATTTTCTTCT	962
Qy	1332	CACAAAACCCACTAAATCTCTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCCCTGCC	1391
Db	963	CAACAAAACCCACTAAATCTCTATCCCCAGAA'TACTGCTGGGATTTATCATATAGGCATGTC	1022
Qy	1392	TTCAGATATTTAAAAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATATTTTGGTTAGAAA	1451
Db	1023	TGTGGATATTTAGGATTTGTCNAGATACCTTGGCAGAAAAAAGAGTATAATTTTGGTTAGAAA	1082
Qy	1452	TAATGTCCTGA	1461
Db	1083	TAACATCTGA	1092

Search completed: May 17, 2002, 19:27:30  
Job time: 12075 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 16:14:02 ; Search time 70.18 Seconds  
(without alignments)  
5250.075 Million cell updates/sec

Title: US-09-863-475A-3

Perfect score: 1500

Sequence: 1 CTTGCCCTTGAGACTTTC.....CTTGACACTATTCTAACCA 1500

Scoring Table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued\_Patents\_NA:\*
- 1: /cgn2\_6/ptodata/2/lna/5A\_COMB.seq:\*
  - 2: /cgn2\_6/ptodata/2/lna/5B\_COMB.seq:\*
  - 3: /cgn2\_6/ptodata/2/lna/6A\_COMB.seq:\*
  - 4: /cgn2\_6/ptodata/2/lna/6B\_COMB.seq:\*
  - 5: /cgn2\_6/ptodata/2/lna/PCTUS\_COMB.seq:\*
  - 6: /cgn2\_6/ptodata/2/lna/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1500	100.0	1500	1	US-07-914-281-3
2	1500	100.0	1500	1	US-08-393-246-3
3	1500	100.0	1500	1	US-08-525-058A-3
4	1500	100.0	1500	2	US-08-696-731-3
5	1500	100.0	1500	4	US-09-042-531-3
6	1454.6	97.0	1500	5	PCT-US91-00899-2
7	1408.4	93.9	3450	2	US-08-378-617A-9
8	1005.8	67.1	1131	2	US-08-704-548-3
9	726.6	48.4	1423	1	US-08-214-580A-2
10	726.6	48.4	1423	5	PCT-US95-07554-5
11	702.8	46.9	1140	2	US-08-704-548-1
12	684.2	45.6	1269	3	US-08-621-700-1
13	684.2	45.6	1269	5	PCT-US95-03940-1
14	684.2	45.6	1412	2	US-08-378-617A-7
15	642.4	42.8	1828	2	US-08-378-617A-8
16	583.2	38.9	1353	1	US-08-214-580A-1
17	541.6	16.1	826	4	US-08-853-774-1
18	240	16.0	1062	1	US-07-752-101A-34
19	233.6	15.6	826	4	US-08-853-774-2
20	232	15.5	826	4	US-08-853-774-3
21	232	15.5	826	4	US-08-853-774-4
22	188.6	12.6	585	1	US-07-752-101A-66
23	83.4	5.6	463	1	US-07-752-101A-67
24	54.6	3.6	7218	1	US-08-232-463-14
25	36.2	2.4	2469	1	US-07-997-133-2
26	36.2	2.4	2469	1	US-08-459-296-1
27	36.2	2.4	2469	5	US-07-997-133-2

c

28	36.2	2.4	2662	2	US-08-451-822A-14	Sequence 14, Appl
29	36.2	2.4	2662	4	US-08-323-430-14	Sequence 14, Appl
30	36.2	2.4	2733	1	US-08-371-001-14	Sequence 14, Appl
31	36.2	2.4	2733	5	PCT-US96-00331-14	Sequence 14, Appl
32	35.8	2.4	282	2	US-08-716-942-18	Sequence 18, Appl
33	35.2	2.3	289	4	US-09-007-005-17	Sequence 17, Appl
34	35.2	2.3	289	4	US-09-244-796-17	Sequence 17, Appl
35	35.2	2.3	547	3	US-09-188-930-14	Sequence 14, Appl
36	34.8	2.3	1416	3	US-08-911-853-3	Sequence 3, Appl1
37	34.8	2.3	1416	4	US-09-479-409-3	Sequence 3, Appl1
38	34.8	2.3	1416	4	US-09-479-453-3	Sequence 3, Appl1
39	34.8	2.3	4377	3	US-08-911-853-28	Sequence 28, Appl
40	34.8	2.3	4377	4	US-09-479-409-28	Sequence 28, Appl
41	34.8	2.3	4377	4	US-09-479-453-28	Sequence 28, Appl
42	34.2	2.3	360	5	PCT-US93-05703-1	Sequence 1, Appl1
43	33.6	2.2	1001	3	US-08-705-771-6	Sequence 6, Appl1
44	33.4	2.2	1179	2	US-08-465-794-4	Sequence 4, Appl1
45	33.4	2.2	1179	3	US-09-049-813-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1  
US-07-914-281-3  
; Sequence 3, Application US/07914281  
; Patent No. 5324663  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/914,281  
; FILING DATE: 19920720  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1500 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; ANTI-SENSE: NO  
; US-07-914-281-3

Query Match 100.0%; Score 1500; DB 1; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTCCCTTGTAGACTCTCTTGAAGTGAAGTACCGATTCTCTGCTGAAGACCTCGCGCT 60  
Db 1 CTTCCCTTGTAGACTCTCTTGAAGTGAAGTACCGATTCTCTGCTGAAGACCTCGCGCT 60  
Qy 61 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCCTTCCCTGCTGAGCCCTCGCTCTT 120  
Db 61 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCCTTCCCTGCTGAGCCCTCGCTCTT 120  
Qy 121 AGGAGCCAGAGCTCGACAGAACTCCGTTGCTGTTGCTGTTGGAGGGAACACAG 180  
Db 121 AGGAGCCAGAGCTCGACAGAACTCCGTTGCTGTTGCTGTTGGAGGGAACACAG 180  
Qy 181 CTGACGATGAGGCTGACTTTGAACCTCAAGAGATCTGTTACCCCAAGTCTCTCGGAATTAA 240  
Db 181 CTGACGATGAGGCTGACTTTGAACCTCAAGAGATCTGTTACCCCAAGTCTCTCGGAATTAA 240  
Qy 241 AGGCTGTACTACATTTGCTGGAACCTCAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Db 241 AGGCTGTACTACATTTGCTGGAACCTCAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Qy 301 CATCTCAACAGATCTCCATGTCAGAGTCCAAAGTCAGAAACAAGTCTTCCATCCTCAAGA 360  
Db 301 CATCTCAACAGATCTCCATGTCAGAGTCCAAAGTCAGAAACAAGTCTTCCATCCTCAAGA 360  
Qy 361 TCTCGATCACAGGAGAAATTAATGAATGTCAGGGAAGAAATCCTGTTGATGCTGATT 420  
Db 361 TCTCGATCACAGGAGAAATTAATGAATGTCAGGGAAGAAATCCTGTTGATGCTGATT 420  
Qy 421 GTCTCAACCTGTTGCTGCTTTTGGGAATATGTCACAGAAATCCAGAGTTGGTGAG 480  
Db 421 GTCTCAACCTGTTGCTGCTTTTGGGAATATGTCACAGAAATCCAGAGTTGGTGAG 480  
Qy 481 AACAGATGGCAGAGGACTGTGCTGCCAAGCTGTTTAAATAATGGACCCACAGTTAT 540  
Db 481 AACAGATGGCAGAGGACTGTGCTGCCAAGCTGTTTAAATAATGGACCCACAGTTAT 540  
Qy 541 CAAGAAGACAACTGTAAGGACGAGAGAAAGGTAAGAAATGGAGATCGCATTTGAAGAG 600  
Db 541 CAAGAAGACAACTGTAAGGACGAGAGAAAGGTAAGAAATGGAGATCGCATTTGAAGAG 600  
Qy 601 CCTAGCTATGGGACTGTTCAATCCAAAGAACCCCGGATGTTTGCACAGTCAACCCG 660  
Db 601 CCTAGCTATGGGACTGTTCAATCCAAAGAACCCCGGATGTTTGCACAGTCAACCCG 660  
Qy 661 TGAAGGCGCGATGTTGTTGGAAGGCACTTATGACACAGCTCTGCTGGAAGTACTAC 720  
Db 661 TGAAGGCGCGATGTTGTTGGAAGGCACTTATGACACAGCTCTGCTGGAAGTACTAC 720  
Qy 721 GCCACAGAAATCACTGTGGGCTGACAGTGTGTTGCTGTTGGGAAAGTACATTGAGCAT 780  
Db 721 GCCACAGAAATCACTGTGGGCTGACAGTGTGTTGCTGTTGGGAAAGTACATTGAGCAT 780  
Qy 781 TACTTTAGAAGACTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTCATA 840  
Db 781 TACTTTAGAAGACTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTCATA 840  
Qy 841 TTTTACGTCATGATAGACGACACCTCCCGGATGCTGCTGTCACCTGAACCCCTCTACAT 900  
Db 841 TTTTACGTCATGATAGACGACACCTCCCGGATGCTGCTGTCACCTGAACCCCTCTACAT 900  
Qy 901 TCCCTTACAAGTCTTTGAGATCAGTCTGAGAAAGGTTGGCAGGATATCAGCATGATCGC 960  
Db 901 TCCCTTACAAGTCTTTGAGATCAGTCTGAGAAAGGTTGGCAGGATATCAGCATGATCGC 960  
Qy 961 ATGAAGACCATTTGGGAGCAGATCCTGCCCCACATCCAGACAGAGTTCGACTTCTTC 1020  
Db 961 ATGAAGACCATTTGGGAGCAGATCCTGCCCCACATCCAGACAGAGTTCGACTTCTTC 1020  
Qy 1021 TGCATGGACGTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAAACTCTGGGCGACGCTG 1080  
Db 1021 TGCATGGACGTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAAACTCTGGGCGACGCTG 1080

## RESULT 2

US-08-393-246-3  
; Sequence 3, Application us/08393246  
; Patent No. 5595900  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,246  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/220,433  
; FILING DATE: 30-MAR-1994  
; APPLICATION NUMBER: US 07/914,281  
; FILING DATE: 20-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1500 base pairs

TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
ANTI-SENSE: NO  
US-09-393-246-3

Query Match 100.0%; Score 1500; DB 1; Length 1500;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CCCTCCCTTGTAGACTCTTCTGGAATGAGAAGTACCGGATTCCTGGAAGACCTCGCGCT	60
DB	1	CCCTCCCTTGTAGACTCTTCTGGAATGAGAAGTACCGGATTCCTGGAAGACCTCGCGCT	60
QY	61	CTCAGGCTCTGGAGTTGGAAACCTGTACCTTCTTCCCTGCTGAGCCCTGCCTCCTT	120
DB	61	CTCAGGCTCTGGAGTTGGAAACCTGTACCTTCTTCCCTGCTGAGCCCTGCCTCCTT	120
QY	121	AGCAGGCCAGAGCTCGACAACTCGGTTGCTTGTCTTGTCTTGGAGGGAACACAG	180
DB	121	AGCAGGCCAGAGCTCGACAACTCGGTTGCTTGTCTTGGAGGGAACACAG	180
QY	181	CTGACGATGAGGCTGACTTTTGAATCAAGAGATCTGCTTACCCAGTCTCCTGGAATTA	240
DB	181	CTGACGATGAGGCTGACTTTTGAATCAAGAGATCTGCTTACCCAGTCTCCTGGAATTA	240
QY	241	AGCCCTGACTACTATTTGCCCTGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC	300
DB	241	AGCCCTGACTACTATTTGCCCTGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC	300
QY	301	CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAGAAACAAGTCTTCCATCCTCAAGA	360
DB	301	CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAGAAACAAGTCTTCCATCCTCAAGA	360
QY	361	TCTGGATCAGAGAGAAATTAATGAATGTCAAGGGAAGTAATCTCTGTGATCTGAT	420
DB	361	TCTGGATCAGAGAGAAATTAATGAATGTCAAGGGAAGTAATCTCTGTGATCTGAT	420
QY	421	GTCTCAACCTGTGTGCTGTTTGGGAATATGTCAACAGATTCACAGAGTTGGTGAG	480
DB	421	GTCTCAACCTGTGTGCTGTTTGGGAATATGTCAACAGATTCACAGAGTTGGTGAG	480
QY	481	AACAGATGGCAGAGGACTGGTGTCCCAAGCTGGTTTAAAAATGGACCCACAGTTAT	540
DB	481	AACAGATGGCAGAGGACTGGTGTCCCAAGCTGGTTTAAAAATGGACCCACAGTTAT	540
QY	541	CAAGAACAACGTAGAGGACGAGAGAAAGGTAGAAATGGAGATCGCATTTGAAGAG	600
DB	541	CAAGAACAACGTAGAGGACGAGAGAAAGGTAGAAATGGAGATCGCATTTGAAGAG	600
QY	601	CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCGCCCGATGTTTGGACAGTACCCCG	660
DB	601	CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCGCCCGATGTTTGGACAGTACCCCG	660
QY	661	TGGAAGCGCGGATGTGTGGAGGACACTTATGACACAGCTCTGCTGGAAAAAGTACTAC	720
DB	661	TGGAAGCGCGGATGTGTGGAGGACACTTATGACACAGCTCTGCTGGAAAAAGTACTAC	720
QY	721	GCCACAGAAACTACTGTGGGCTGACAGTGTCTGCTGGGAAAGTACATTTGAGCAT	780
DB	721	GCCACAGAAACTACTGTGGGCTGACAGTGTCTGCTGGGAAAGTACATTTGAGCAT	780
QY	781	TACTTAGAAGACTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTATA	840
DB	781	TACTTAGAAGACTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTATA	840
QY	841	TTTTACCTCATATGATAGACACACCTCCCGGATGCTCTGCTGCACTTGAACCTCTACAT	900
DB	841	TTTTACCTCATATGATAGACACACCTCCCGGATGCTCTGCTGCACTTGAACCTCTACAT	900
QY	901	TCCTTACAAGTCTTTGAGATCAGGCTGTGAGAAGAGGTGGCAGGATATCAGCATGATCGC	960

DB	901	TCCTTACAAGTCTTTGAGATCAGGCTGTGAGAAGAGGTGGCAGGATATCAGCATGATCGC	960
QY	961	ATGAAGACCATTCGGGAGCAGCATCTCTGGCCACATCCAGCACGAGTCTGACTTCTCTTC	1020
DB	961	ATGAAGACCATTCGGGAGCAGCATCTCTGGCCACATCCAGCACGAGTCTGACTTCTCTTC	1020
QY	1021	TGCATGGACCTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAAACTCTGGGCCAGCTG	1080
DB	1021	TGCATGGACCTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAAACTCTGGGCCAGCTG	1080
QY	1081	GTAGCAGACTTCAGAGCCCTGGTGTACAGAGCCAGTCCCGAGAGTTCACCTATGAGAGG	1140
DB	1081	GTAGCAGACTTCAGAGCCCTGGTGTACAGAGCCAGTCCCGAGAGTTCACCTATGAGAGG	1140
QY	1141	CGGAAGTCTGCGCGGCTACATTTCCATTCGGAGAGGGGATTTTACTACACGCGGCC	1200
DB	1141	CGGAAGTCTGCGCGGCTACATTTCCATTCGGAGAGGGGATTTTACTACACGCGGCC	1200
QY	1201	ATTTTGGAGAACGCTTACTCACAATTCCTCAACCTTCACGAGGAGTCTTTAAGGGATC	1260
DB	1201	ATTTTGGAGAACGCTTACTCACAATTCCTCAACCTTCACGAGGAGTCTTTAAGGGATC	1260
QY	1261	CTCAGGACAAGAAACATGACATAGAACCCAGTGGCATGATGAGGCCACCTCAACAAA	1320
DB	1261	CTCAGGACAAGAAACATGACATAGAACCCAGTGGCATGATGAGGCCACCTCAACAAA	1320
QY	1321	TACTTCTCTTTCAACAAACCCACTTAAATCTTCTCAGAGTATTTCTGGGACTATCAG	1380
DB	1321	TACTTCTCTTTCAACAAACCCACTTAAATCTTCTCAGAGTATTTCTGGGACTATCAG	1380
QY	1381	ATAGGCTGCTTCAGATATTAAGTGTCAAGTGTCTGGCAGACAAAGAGTATAT	1440
DB	1381	ATAGGCTGCTTCAGATATTAAGTGTCAAGTGTCTGGCAGACAAAGAGTATAT	1440
QY	1441	TTGGTTAGAAATATGTCTGACTTCAATTTGTGATGGAACCTTGACACTATTTCTAACCA	1500
DB	1441	TTGGTTAGAAATATGTCTGACTTCAATTTGTGATGGAACCTTGACACTATTTCTAACCA	1500

RESULT 3

US-08-525-058A-3

; Sequence 3, Application US/08525058A

; Patent No. 5770420

; GENERAL INFORMATION:

; APPLICANT: LOWE, JOHN B.

; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS

; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.

; STREET: 1755 Jefferson Davis Highway, Fourth Floor

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/525,058A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Lavalleye, Jean-Paul M. P.

; REGISTRATION NUMBER: 31,451

; REFERENCE/DOCKET NUMBER: 2363-060-55

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)521-4500

TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1500 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ANTI-SENSE: NO  
US-08-525-058A-3

Query Match 100.0%; Score 1500; DB 1; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCCCTTGTAGACTCTTCTTGAAGTGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Db 1 CCTTCCCTTGTAGACTCTTCTTGAAGTGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Qy 61 CTCAGGCTCTGGAGTGTGAACCTGTACCTTCTTCTCTGCTGAGCCCTGCTCCTT 120  
Db 61 CTCAGGCTCTGGAGTGTGAACCTGTACCTTCTTCTCTGCTGAGCCCTGCTCCTT 120  
Qy 121 AGGAGGCCAGAGCTCGACAGAACTCGGTGCTTGTGTTTGGTTTGGAGGAACACAG 180  
Db 121 AGGAGGCCAGAGCTCGACAGAACTCGGTGCTTGTGTTTGGTTTGGAGGAACACAG 180  
Qy 181 CTGAGGATGAGGCTGACTTGAACCTCAAGAGATCTGCTTACCCAGTCTCTGGAATTA 240  
Db 181 CTGAGGATGAGGCTGACTTGAACCTCAAGAGATCTGCTTACCCAGTCTCTGGAATTA 240  
Qy 241 AGGCTGTACTACATTTGCTTGGACCTAAGATTTTCATGATCACTATGCTTCAAGATCT 300  
Db 241 AGGCTGTACTACATTTGCTTGGACCTAAGATTTTCATGATCACTATGCTTCAAGATCT 300  
Qy 301 CATGTCAACAAGATCTCCATGTCAGATCCAAAGTCCAGAACAAAGTCTTCCATCTCAAGA 360  
Db 301 CATGTCAACAAGATCTCCATGTCAGATCCAAAGTCCAGAACAAAGTCTTCCATCTCAAGA 360  
Qy 361 TCTGGATCACAGGAGAAATTAATGAATGTCAGGGAAGAAATGCTTCTGCTGATGCTGATT 420  
Db 361 TCTGGATCACAGGAGAAATTAATGAATGTCAGGGAAGAAATGCTTCTGCTGATGCTGATT 420  
Qy 421 GTCTCAACCGTGTGTGCTGTTTGGGAATATGTCACAGAAATTCACAGAGTTGGTGAG 480  
Db 421 GTCTCAACCGTGTGTGCTGTTTGGGAATATGTCACAGAAATTCACAGAGTTGGTGAG 480  
Qy 481 AACAGATGGCAGAGGACTGTGTTTCCCAAGCTGGTTTAAATGGGACCCACAGTTAT 540  
Db 481 AACAGATGGCAGAGGACTGTGTTTCCCAAGCTGGTTTAAATGGGACCCACAGTTAT 540  
Qy 541 CAAGAACAACAGTGAAGAGGAGGAGAAAGGTAGAAATGGAGATCGCATTTGAAGAG 600  
Db 541 CAAGAACAACAGTGAAGAGGAGGAGAAAGGTAGAAATGGAGATCGCATTTGAAGAG 600  
Qy 601 CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGGATGTTTGGACAGTACCCCG 660  
Db 601 CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGGATGTTTGGACAGTACCCCG 660  
Qy 661 TGGAGGCGCGATGTTGGGAGGCACTTATGACACAGCTCTGCTGGGAAAGTACTAC 720  
Db 661 TGGAGGCGCGATGTTGGGAGGCACTTATGACACAGCTCTGCTGGGAAAGTACTAC 720  
Qy 721 GCCACAGAAGTCACTGTGGGCTGACAGTGTGCTGTGGGAAAGTACATGAGCAT 780  
Db 721 GCCACAGAAGTCACTGTGGGCTGACAGTGTGCTGTGGGAAAGTACATGAGCAT 780  
Qy 781 TACTTAGAAGACTTTCTGGAGTCTGCTGACATGACTTTTCATGTTGGCCATCGGGTCATA 840  
Db 781 TACTTAGAAGACTTTCTGGAGTCTGCTGACATGACTTTTCATGTTGGCCATCGGGTCATA 840

Qy 841 TTTTACGTCATGATAGACGACACCTCCCGGATGCTGTGTCACCTGAACCCCTCTACAT 900  
Db 841 TTTTACGTCATGATAGACGACACCTCCCGGATGCTGTGTCACCTGAACCCCTCTACAT 900  
Qy 901 TCCTTACAAGTCTTTGAGATCAGTCTCAGAAGAGGTGGCAGGATATCAGCATGATGCGC 960  
Db 901 TCCTTACAAGTCTTTGAGATCAGTCTCAGAAGAGGTGGCAGGATATCAGCATGATGCGC 960  
Qy 961 ATGAAGACCATTTGGGAGCACATCTGCGCCACATCCAGCAGAGTGCAGTCTCTCTTC 1020  
Db 961 ATGAAGACCATTTGGGAGCACATCTGCGCCACATCCAGCAGAGTGCAGTCTCTCTTC 1020  
Qy 1021 TGATGAGGATGATCAAGTCTTTTCAAGACAACTTTCGGGGTGAAGTCTTGGGCCAGT 1080  
Db 1021 TGATGAGGATGATCAAGTCTTTTCAAGACAACTTTCGGGGTGAAGTCTTGGGCCAGT 1080  
Qy 1081 GTAGCACAGCTCCAGGCTGCTGTGTACAAAGCCAGTCCCGAGAGTTTCACTATGAGAG 1140  
Db 1081 GTAGCACAGCTCCAGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140  
Qy 1141 CGGGAACCTGTCGGCGCGTACATTTCCATTTCGAGAGAGGGGATTTTACTACACGCGCC 1200  
Db 1141 CGGGAACCTGTCGGCGCGTACATTTCCATTTCGAGAGAGGGGATTTTACTACACGCGCC 1200  
Qy 1201 ATTTTGGAGGAGCGCTTACTACATTTCTCAACCTTCCAGGAGTGTCTTAAAGGGATC 1260  
Db 1201 ATTTTGGAGGAGCGCTTACTACATTTCTCAACCTTCCAGGAGTGTCTTAAAGGGATC 1260  
Qy 1261 CTCAGGACAAAGAACATGACATAGAGCCAGTGGCATGATGAGAGCCACTCAACAA 1320  
Db 1261 CTCAGGACAAAGAACATGACATAGAGCCAGTGGCATGATGAGAGCCACTCAACAA 1320  
Qy 1321 TACTTCTTTTCAACAAACCCACTTCTATCTCCAGAGTATTTGCTGGGACTATCAG 1380  
Db 1321 TACTTCTTTTCAACAAACCCACTTCTATCTCCAGAGTATTTGCTGGGACTATCAG 1380  
Qy 1381 ATAGCCTGCTTCCAGATATTTAAAGTGTCAAGGTAGTGTGGCAGACAAAGATTAAT 1440  
Db 1381 ATAGCCTGCTTCCAGATATTTAAAGTGTCAAGGTAGTGTGGCAGACAAAGATTAAT 1440  
Qy 1441 TTGGTTAGAAATATGCTGACTTCAAAATGCTGATGGAACCTTGAACCTATTTCTAACCA 1500  
Db 1441 TTGGTTAGAAATATGCTGACTTCAAAATGCTGATGGAACCTTGAACCTATTTCTAACCA 1500

RESULT 4  
US-08-696-731-3  
Sequence 3, Application US/08696731  
Patent No. 5955347  
GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/696.731  
FILING DATE: 14-AUG-1996  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/393,246  
FILING DATE:  
APPLICATION NUMBER: US 08/220,433  
FILING DATE: 30-MAR-1994  
APPLICATION NUMBER: US 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1500 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
ANTI-SENSE: NO  
US-08-696-731-3

Query Match 100.0%; Score 1500; DB 2; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTTCCCTTTAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAACACCTCGCGCT 60  
DB 1 CCTTCCCTTTAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAACACCTCGCGCT 60  
QY 61 CTGAGGCTCGGAGTGGAACTCTGACCTCTCTTCCCTCTGCTGAGCCCTGCCTCTT 120  
DB 61 CTGAGGCTCGGAGTGGAACTCTGACCTCTCTTCCCTCTGCTGAGCCCTGCCTCTT 120  
QY 121 AGGAGCCAGAGCTCGACAGAACTCGGTTGCTTGTGCTTGTGCTTGGAGGAAACACAG 180  
DB 121 AGGAGCCAGAGCTCGACAGAACTCGGTTGCTTGTGCTTGTGCTTGGAGGAAACACAG 180  
QY 181 CTGAGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCCTGGAATTAA 240  
DB 181 CTGAGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCCTGGAATTAA 240  
QY 241 AGGCTCTACTACATTTGCTGAGCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
DB 241 AGGCTCTACTACATTTGCTGAGCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
QY 301 CATGTCAACAGATCTCCATGTCAAGATCCAAGTCAAGAACTCAAGTCTTCCATCCTCAAGA 360  
DB 301 CATGTCAACAGATCTCCATGTCAAGATCCAAGTCAAGAACTCAAGTCTTCCATCCTCAAGA 360  
QY 361 TCTGGATCAGAGAGAAATTAATGATGTCAAGGAGAAAGTAACTCTGTTGATGCTGATT 420  
DB 361 TCTGGATCAGAGAGAAATTAATGATGTCAAGGAGAAAGTAACTCTGTTGATGCTGATT 420  
QY 421 GTCTCAACCTGTTGCTGTTTGGGAATATGTCAACAGAAATTCACAGAGTTGGTGAG 480  
DB 421 GTCTCAACCTGTTGCTGTTTGGGAATATGTCAACAGAAATTCACAGAGTTGGTGAG 480  
QY 481 AACAGATGGCAGAGGACTGGTGTCTCCCAAGCTGGTTTAAATGGGACCCACAGTTAT 540  
DB 481 AACAGATGGCAGAGGACTGGTGTCTCCCAAGCTGGTTTAAATGGGACCCACAGTTAT 540  
QY 541 CAGAAGACAACTAGAGGACGAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAG 600  
DB 541 CAGAAGACAACTAGAGGACGAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAG 600  
QY 601 CCTCAGCTATGGAGTGGTTCAATCCAAAGACCGCCCGATGTTTTCAGAGTGCACCCG 660  
DB 601 CCTCAGCTATGGAGTGGTTCAATCCAAAGACCGCCCGATGTTTTCAGAGTGCACCCG 660

QY 661 TGAAGGCGCGGATTTGTGGGAAGGACACTTATGACACAGCTCTGCTGGAAAAAGTACTAC 720  
DB 661 TGAAGGCGCGGATTTGTGGGAAGGACACTTATGACACAGCTCTGCTGGAAAAAGTACTAC 720  
QY 721 GCCACAGAAACTCACTGTGGGGCTGACAGTGTCTGCTGGGAAAGTACATTTGAGCAT 780  
DB 721 GCCACAGAAACTCACTGTGGGGCTGACAGTGTCTGCTGGGAAAGTACATTTGAGCAT 780  
QY 781 TACTTGAAGACACTTTCTGGAGTCTGCTGACATGTACTTCAATGTTGGCCATCGGTCATA 840  
DB 781 TACTTGAAGACACTTTCTGGAGTCTGCTGACATGTACTTCAATGTTGGCCATCGGTCATA 840  
QY 841 TTTTACGTCATGATAGACACACTCTCCGGATGCTGCTGCACTTGAACCTCTTACAT 900  
DB 841 TTTTACGTCATGATAGACACACTCTCCGGATGCTGCTGCACTTGAACCTCTTACAT 900  
QY 901 TCCTTACAAGTCTTTGAGATCAGTCTGAGAAGAGTGGGAGGATATCAGATGATCGC 960  
DB 901 TCCTTACAAGTCTTTGAGATCAGTCTGAGAAGAGTGGGAGGATATCAGATGATCGC 960  
QY 961 ATGAAGACCAATTGGGAGACACATCTCGGCCACATCCAGCAGAGTCTGACTTCTCTTC 1020  
DB 961 ATGAAGACCAATTGGGAGACACATCTCGGCCACATCCAGCAGAGTCTGACTTCTCTTC 1020  
QY 1021 TGCATGGACGCTGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAACTCTGGGCCAGCTG 1080  
DB 1021 TGCATGGACGCTGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAACTCTGGGCCAGCTG 1080  
QY 1081 GTAGCAGAGCTCCAGGCTGCTGTAAGAGGAGTCCCGAGAGTTCACCTATGAGAGG 1140  
DB 1081 GTAGCAGAGCTCCAGGCTGCTGTAAGAGGAGTCCCGAGAGTTCACCTATGAGAGG 1140  
QY 1141 CGGAACTGTCGGCGGTACATTCATTCGAGAGGAGGATTTTACTACCGCGGCC 1200  
DB 1141 CGGAACTGTCGGCGGTACATTCATTCGAGAGGAGGATTTTACTACCGCGGCC 1200  
QY 1201 ATTTTGGAGAACGCTACTCATCTTCAACCTCAACGAGGAGTCTTTAAGGGATC 1260  
DB 1201 ATTTTGGAGAACGCTACTCATCTTCAACCTCAACGAGGAGTCTTTAAGGGATC 1260  
QY 1261 CTCAGGACAGAAACATGACATAGAGCCAGTGGCATGATGAGAGCCACCTCAACAA 1320  
DB 1261 CTCAGGACAGAAACATGACATAGAGCCAGTGGCATGATGAGAGCCACCTCAACAA 1320  
QY 1321 TACTTCTTTTCAACAAACCCACTTAAATCTTCTCAGAGTATTTCTGGGACTATCAG 1380  
DB 1321 TACTTCTTTTCAACAAACCCACTTAAATCTTCTCAGAGTATTTCTGGGACTATCAG 1380  
QY 1381 ATAGGCTGCTTTCAGATATTTAAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCA 1440  
DB 1381 ATAGGCTGCTTTCAGATATTTAAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCA 1440  
QY 1441 TTGGTTAGAATAATGCTGCTCAATTTGTGATGAAACTTGACACTATTTCTACCA 1500  
DB 1441 TTGGTTAGAATAATGCTGCTCAATTTGTGATGAAACTTGACACTATTTCTACCA 1500

RESULT 5  
US-09-042-531-3  
Sequence 3, Application US/09042531  
Patent No. 6268193  
GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
RES  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/042,531  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,246  
FILING DATE:  
APPLICATION NUMBER: US 08/220,433  
FILING DATE: 30-MAR-1994  
APPLICATION NUMBER: US 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1500 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
ANTI-SENSE: NO

US-09-042-531-3

Query Match 100.0%; Score 1500; DB 4; Length 1500;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCCCTGTAGACTCTTCTTGGAAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Db 1 CCTTCCCTGTAGACTCTTCTTGGAAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Qy 61 CTCAGGCTCTGGAGTTGGAAACCTGTACCTTCTTCTCTGCTGAGCCCTGCTT 120  
Db 61 CTCAGGCTCTGGAGTTGGAAACCTGTACCTTCTTCTCTGCTGAGCCCTGCTT 120  
Qy 121 AGGCAGGCGCAGAGCTCGACAGAACTCGGTTGCTTGTGCTTGTGGAGGGAACACAG 180  
Db 121 AGGCAGGCGCAGAGCTCGACAGAACTCGGTTGCTTGTGCTTGTGGAGGGAACACAG 180  
Qy 181 CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCTGGAATTA 240  
Db 181 CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCTGGAATTA 240  
Qy 241 AGGCTGTACTACATTTGCTTGGACCTGAAGATTTTCATGATCACTATGCTTCAAGATCT 300  
Db 241 AGGCTGTACTACATTTGCTTGGACCTGAAGATTTTCATGATCACTATGCTTCAAGATCT 300  
Qy 301 CATGTCAACAAGATCTCCATGTCGAAGTCCAAAGTCAGAAACAAAGTCTTCCATCTCAAGA 360  
Db 301 CATGTCAACAAGATCTCCATGTCGAAGTCCAAAGTCAGAAACAAAGTCTTCCATCTCAAGA 360  
Qy 361 TCTGGATCAGAGGAGAAATTAATGAATGTCGAAGGAAAGTAATCTGTTGATGCTGATT 420  
Db 361 TCTGGATCAGAGGAGAAATTAATGAATGTCGAAGGAAAGTAATCTGTTGATGCTGATT 420  
Qy 421 CTCCTCAACCGTGTGCTGCTTTGGGAATATGTCAACAGAAATTCACAGAGTTGGTGAG 480  
Db 421 CTCCTCAACCGTGTGCTGCTTTGGGAATATGTCAACAGAAATTCACAGAGTTGGTGAG 480

...

Sequence 2, Application PC/TUS9100899  
GENERAL INFORMATION:  
APPLICANT: Lowe, John B.  
TITLE OF INVENTION: Method and Products For the Synthesis of  
TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,  
TITLE OF INVENTION: or as Free Molecules, and For the Isolation of Cloned  
TITLE OF INVENTION: Genetic Sequences That Determine These Structures  
NUMBER OF INVENTIONS: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/00899  
FILING DATE: 19910214  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye Ph.D., Jean-Paul  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-021-55 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-5940  
TELEFAX: (703)486-2347  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1500 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cdna  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: Blood  
TISSUE TYPE: Blood  
PCT-US91-00899-2

Query Match 97.0%; Score 1454.6; DB 5; Length 1500;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1485; Conservative 0; Mismatches 14; Indels 2; Gaps 2:

QY 1 CCTTCCCTGTAGACTCTTCTTGGAAATGAGAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
DB 1 CCTTCCCTGTAGACTCTTCTTGGAAATGAGAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
QY 61 CTCAGGCTCTGGAGTTGGAACCTCTACTTCTTCTGCTGCTGAGCCCTGCTTCTTCT 120  
DB 61 CTCAGGCTCTGGAGTTGGAACCTCTACTTCTTCTGCTGCTGAGCCCTGCTTCTTCT 120  
QY 121 AGCAGGCCAGAGCTCGACAGACTCGGTGCTTCTGCTTCTGCTTGGAGGGAACACAG 180  
DB 121 AGCAGGCCAGAGCTCGACAGACTCGGTGCTTCTGCTTCTGCTTGGAGGGAACACAG 180  
QY 181 CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCTCTGGAATTA 240  
DB 181 CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCTCTGGAATTA 240  
QY 241 AGCCCTGTACTACTTTGCTTGGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
DB 241 AGCCCTGTACTACTTTGCTTGGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
QY 301 CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAGAACAGTCTTCCATCCTCAAGA 360  
DB 301 CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAGAACAGTCTTCCATCCTCAAGA 360

QY 361 TCTGGATCACAGGAGAAAAATAATGTCAAGGGAAAAAGTAATCCTGTGTGATGCTGATT 420  
DB 361 TCTGGATCACAGGAGAAAAATAATGTCAAGGGAAAAAGTAATCCTGTGTGATGCTGATT 420  
QY 421 GTCTCAACCGTGGTTGCTGCTGTTTGGGAATATGTCAACAGAAATTCACAGGTTGGTGAG 480  
DB 421 GTCTCAACCGTGGTTGCTGCTGTTTGGGAATATGTCAACAGAAATTCACAGGTTGGTGAG 480  
QY 481 AACAGATGCCAGAAGACTGGTGGTTCCCAAGCTGTTTAAAAAATGGAGCCACAGTTAT 540  
DB 481 AACAGATGCCAGAAGACTGGTGGTTCCCAAGCTGTTTAAAAAATGGAGCCACAGTTAT 540  
QY 541 CAAGAAGACAACTAGAAAGGACGAGAGAAAGGGTAGAAATGGAGATCGCATTTGAAGAG 600  
DB 541 CAAGAAGACAACTAGAAAGGACGAGAGAAAGGGTAGAAATGGAGATCGCATTTGAAGAG 600  
QY 601 CCTCAGCTATGGACTGGTTCAATCCAAAGAACCCCGCCGGATGTTTGTACAGTGAACCCG 660  
DB 601 CCTCAGCTATGGACTGGTTCAATCCAAAGAACCCCGCCGGATGTTTGTACAGTGAACCCG 660  
QY 661 TGAAGGCGCGGATGCTGTTGGGAAGGCACTTATGACACAGCTCTGCTGGAAGAGTACTAC 720  
DB 661 TGAAGGCGCGGATGCTGTTGGGAAGGCACTTATGACACAGCTCTGCTGGAAGAGTACTAC 720  
QY 721 GCCACACAGAACTACTGTGGGGCTGACAGTGTGCTGTGGGAAAGTACATTTGAGCAT 780  
DB 721 GCCACACAGAACTACTGTGGGGCTGACAGTGTGCTGTGGGAAAGTACATTTGAGCAT 780  
QY 781 TACTTAAAGACTTTCTGGAGTCTGCTGACATGATCTTCAATGTTGGCCATCGGGTCAATA 840  
DB 781 TACTTAAAGACTTTCTGGAGTCTGCTGACATGATCTTCAATGTTGGCCATCGGGTCAATA 840  
QY 841 TTTTACGTGATGATACGACACCTCCCGGATGCTGCTGACACCTGAAACCTCTACAT 900  
DB 841 TTTTACGTGATGATACGACACCTCCCGGATGCTGCTGACACCTGAAACCTCTACAT 900  
QY 901 TCCTTAAAGTCTTTGAGATCAGTCTGAGAGAGGTGGCAGGATATCAGCATGATGCGC 960  
DB 901 TCCTTAAAGTCTTTGAGATCAGTCTGAGAGAGGTGGCAGGATATCAGCATGATGCGC 960  
QY 961 ATGAAGACCATTTGGGGAGCACATCTTGGCCACATCCAGCAGAGGTGAGTCTCTTTC 1020  
DB 961 ATGAAGACCATTTGGGGAGCACATCTTGGCCACATCCAGCAGAGGTGAGTCTCTTTC 1020  
QY 1021 TGCAATGAGTGGATCAAGTCTTTCAAGCAACTTCGGGTGCAAACTCTGGGCCAGCTG 1080  
DB 1021 TGCAATGAGTGGATCAAGTCTTTCAAGCAACTTCGGGTGCAAACTCTGGGCCAGCTG 1080  
QY 1081 GTAGCACAGCTCCAGGCTGGTGTACAAAGGCCAGTCCCGAGAAAGTTTCACTATGAGAG 1140  
DB 1081 GTAGCACAGCTCCAGGCTGGTGTACAAAGGCCAGTCCCGAGAAAGTTTCACTATGAGAG 1140  
QY 1141 CGGGAAGTCTGGGCCGCTGATCTTCCATTCGGAGAGGGGATTTTACTACACAGCGGCC 1200  
DB 1141 CGGGAAGTCTGGGCCGCTGATCTTCCATTCGGAGAGGGGATTTTACTACACAGCGGCC 1200  
QY 1201 ATTTTGGAGGAGCGCTACTACATTTCTCAACCTCACAGGAGTGTCTTAAAGGGAT 1259  
DB 1201 ATTTTGGAGGAGCGCTACTACATTTCTCAACCTCACAGGAGTGTCTTAAAGGGAT 1259  
QY 1260 CTCTCCAGGACAAAGAACATGACATAGAAGCCAGTGGCATGATGAGAGCCACCTCAACAA 1319  
DB 1260 CTCTCCAGGACAAAGAACATGACATAGAAGCCAGTGGCATGATGAGAGCCACCTCAACAA 1319  
QY 1320 ATACTTCTTTTCAACAAACCCACTAAATCTCTATCTCCAGATATTTGCTGGAGCTATCA 1379  
DB 1320 ATACTTCTTTTCAACAAACCCACTAAATCTCTATCTCCAGATATTTGCTGGAGCTATCA 1379  
QY 1380 GATAGGCTCTGCTTCAAGATTTAAAGTGTCAAGGTAGCTTGGCAGACAAAGATATAA 1439  
DB 1380 GATAGGCTCTGCTTCAAGATTTAAAGTGTCAAGGTAGCTTGGCAGACAAAGATATAA 1439

QY 1440 TTTGGTTAGAAATAATGCTGACTTCAAAATGTGATGGAAACCTTGACACACTATTTCCTAAC 1499  
Db 1440 TTTGGTTAGAAATAATGCTGACTTCAAAATGTGATGGAAACCTTGACACACTATTTCCTAAC 1499  
QY 1500 A 1500  
Db 1500 A 1500  
RESULT 7  
US-08-378-617A-9  
; Sequence 9, Application US/08378617A  
; Patent No. 5849991  
; GENERAL INFORMATION:  
; APPLICANT: d'Apice, Anthony J. F.  
; APPLICANT: Pearse, Martin J.  
; APPLICANT: Robins, Allan J.  
; APPLICANT: Crawford, Robert J.  
; APPLICANT: Rathjen, Peter D.  
; TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF  
; TITLE OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 120 South Sixth Street, Suite 2500  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/378,617A  
; FILING DATE: 26-JAN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 06868/005001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (612) 335-5070  
; TELEFAX: (612) 288-9696  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3450 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-378-617A-9  
Query Match 93.9%; Score 1408.4; DB 2; Length 3450;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 1483; Conservative 0; Mismatches 11; Indels 38; Gaps 3;  
QY 1 CCTTCCCTCTGAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTTCGGCT 60  
Db 64 CCTTCCCTCTGAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTTCGGCT 123  
QY 61 CTCAGGCTCTGGGAGTTGGAAACCTTGTACCTTTCCTTTCCTGCTGAGCCCTGCCTCCT 119  
Db 124 CTCAGGCTCTGGGAGTTGGAAACCTTGTACCTTTCCTTTCCTGCTGAGCCCTGCCTCCT 183  
QY 120 TAGCAGGCCAGAGCTCGACAGAACTCGGTTGCTTGTCTGCTTTGGAGGGAACACA 179  
Db 184 TCGCAGGCCAGAGCTGACAGAACTCGGTTGCTTGTCTTGTGCTTTGGAGGGAACACA 243  
QY 180 GCTGACCATGAGCTGACTTTGAATCAAGAGATCTGCTTACCCAGTCTCCTGGAAATTA 239

Db 244 GCTGACGATGAGGCTGACTTTTGAATCAAGAGATCTGCTTACCCAGTCTCCTGGAATTA 303  
QY 240 AAGGCTGTACTATCTTGCCTGGACCTTAAGATTTTTCATGATCACTATGCTTCAAGATCT 299  
Db 304 AAGGCTGTACTAC-CTTGCCTGGACCTTAAGATTTTTCATGATCACTATGCTTCAAGATCT 362  
QY 300 CCATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAACAGTCTTCCATCCTCAAG 359  
Db 363 CCATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAACAGTCTTCCATCCTCAAG 422  
QY 360 ATCTGGATCACAGGAGAAATAATGAATGTCAAGGGGAAAGTAATCCTGTTGATGCTGAT 419  
Db 423 ATCTGGATCACAGGAGAAATAATGAATGTCAAGGGGAAAGTAATCCTGTTGATGCTGAT 482  
QY 420 TGTCTCAACCGTGGTTGCTGCTGTTTGGGAATATGTCAACAG----- 461  
Db 483 TGTCTCAACCGTGGTTGCTGCTGTTTGGGAATATGTCAACAGCCAGACGGCTCTTCTTT 542  
QY 462 -----AATTCAGAGGTTGGTGAGAACAGATGGCAGAGAGACTGGTG 503  
Db 543 GTGGATATATCACACAAAAATTCAGAGGTTGGTGAGAACAGATGGCAGAGAGACTGGTG 602  
QY 504 GTTCCCAAGCTGTTTAAAAATGGGACCCACAGTTTCAAGAAGACAACTAGAAGGACG 563  
Db 603 GTTCCCAAGCTGTTTAAAAATGGGACCCACAGTTTCAAGAAGACAACTAGAAGGACG 662  
QY 564 GAGAGAAAGGTAGAAATGGAGATCGCATTTGAAGGCCTCAGCTATGGGACTGGTTCAA 623  
Db 663 GAGAGAAAGGTAGAAATGGAGATCGCATTTGAAGGCCTCAGCTATGGGACTGGTTCAA 722  
QY 624 TCCAAAGAACCGCCGGATGTTTTCAGAGTACCCCGTGAAGCGCCGATGTTGTGGGA 683  
Db 723 TCCAAAGAACCGCCGGATGTTTTCAGAGTACCCCGTGAAGCGCCGATGTTGTGGGA 782  
QY 684 AGGACTTATGACACAGCTCTGCTGGAAAGTACTACGCCACACAGAACTCACTGTGGG 743  
Db 783 AGGACTTATGACACAGCTCTGCTGGAAAGTACTACGCCACACAGAACTCACTGTGGG 842  
QY 744 GCTGACAGTCTTTGCTGTGGGAAGTACATTGAGCATTTACTTAGAAGACTTTCTGGAGTC 803  
Db 843 GCTGACAGTCTTTGCTGTGGGAAGTACATTGAGCATTTACTTAGAAGACTTTCTGGAGTC 902  
QY 804 TGCTGACATGTACTTTCATGTTGGCCATCGGTCATATTTTACGTATGATAGACGAC 863  
Db 903 TGCTGACATGTACTTTCATGTTGGCCATCGGTCATATTTTACGTATGATAGACGAC 962  
QY 864 CTCCCGGATGCTGCTGTCGACCTGAACCCCTTACATTCCTTACAAGTCTTTGAGATCAG 923  
Db 963 CTCCCGGATGCTGCTGTCGACCTGAACCCCTTACATTCCTTACAAGTCTTTGAGATCAG 1022  
QY 924 GTCTGAGAAGAGTGGCAGGATATCAGCATGTATGCGCATGAAGACCATTTGGGAGCAGAT 983  
Db 1023 GTCTGAGAAGAGTGGCAGGATATCAGCATGTATGCGCATGAAGACCATTTGGGAGCAGAT 1082  
QY 984 CTTGGCCCATCCAGCAGGCTGACATTTCTTCTGCTGATGGACGTGATCAAGTCTTT 1043  
Db 1083 CTTGGCCCATCCAGCAGGCTGACATTTCTTCTGCTGATGGACGTGATCAAGTCTTT 1142  
QY 1044 TCAAGACAACCTTCGGGTGAAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCTCGGTG 1103  
Db 1143 TCAAGACAACCTTCGGGTGAAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCTCGGTG 1202  
QY 1104 GTACAAGCCAGTCCCAGAGAGTTACCTATGAGAGCGGGAACTGTGCGCCGCGTACAT 1163  
Db 1203 GTACAAGCCAGTCCCAGAGAGTTACCTATGAGAGCGGGAACTGTGCGCCGCGTACAT 1262  
QY 1164 TCCATTCGGAGAGGGGATTTTACTACACGCGGGCATTTTGGAGGAGACGCTACTCA 1223  
Db 1263 TCCATTCGGAGAGGGGATTTTACTACACGCGGGCATTTTGGAGGAGACGCTACTCA 1322  
QY 1224 CATTTCTCAACTCACCAGGAGTGTCTTAAAGGGATCCTCCAGGACAAAGAAATGACAT 1283  
Db 1323 CATTTCTCAACTCACCAGGAGTGTCTTAAAGGGATCCTCCAGGACAAAGAAATGACAT 1382



QY 1284 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCAC 1343  
Db 1383 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCAC 1442  
QY 1344 TAAATCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCGTGGCTTCAGATATTAA 1403  
Db 1443 TAAATCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCGTGGCTTCAGATATTAA 1502  
QY 1404 AAGTGTCAAGCTAGCTTGGCAGACAAAGAGTATAATTTGGTTAGAAATATCTCTGACT 1463  
Db 1503 AAGTGTCAAGCTAGCTTGGCAGACAAAGAGTATAATTTGGTTAGAAATATCTCTGACT 1562  
QY 1464 TCAATTTGTGATGGAATTTTGACACTATTCTT 1495  
Db 1563 TCAATTTGTGATGGAATTTTGACACTATTACT 1594

RESULT 8  
US-08-704-548-3  
; Sequence 3, Application US/08704548  
; Patent No. 5879675  
; GENERAL INFORMATION:  
; APPLICANT: GALILI, URI  
; TITLE OF INVENTION: REPIK, PATRICIA M.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR VACCINES  
; TITLE OF INVENTION: COMPRISING ALPHA-GALACTOSYL EPITOPES  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEIDEL, GONDA, LAVORNA & MONACO, P.C.  
; STREET: Suite 1800, Two Penn Center Plaza  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19102

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/704,548  
; FILING DATE: 11-SEP-1996  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 8760-2 CII  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1131 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA

US-08-704-548-3

Query Match 67.1%; Score 1005.8; DB 2; Length 1131;  
Best Local Similarity 95.3%; Pred. No. 3.7e-280;  
Matches 1078; Conservative 0; Mismatches 2; Indels 51; Gaps 2;

QY 382 ATGAATCTCAAGGAAAGTAATCTCTTGTATGCTGATGCTCAACCGGTGGTGTGCGT 441  
Db 1 ATCAATCTCAAGGAAAGTATCTCTTGTATGCTGATGCTCAACCGGTGGTGTGCGT 60  
QY 442 TTTTGGGAATATGTCACACG-----NATT 465  
Db 61 TTTTGGGAATATGTCACACGCGCCACGCGCTCTTCTTGTGGATATATCACAAAAAATT 120

QY 466 CCAGAGTTGGTGGAGAACAGATGGCAGAGGACTGGTGGTTCCCAAGCTGGTTTAAAAAT 525  
Db 121 CCAGAGTTGGTGGAGAACAGATGGCAGAGGACTGGTGGTTCCCAAGCTGGTTTAAAAAT 180  
QY 526 GGACCCACAGTTTATCAAGAGACACAGTAGAAGGACGGAGAGAAAAGGCTAGAAATGGA 585  
Db 181 GGACCCACAGTTTATCAAGAGACACAGTAGAAGGACGGAGAGAAAAGGCTAGAAATGGA 240  
QY 586 GATCGCATTTGA-----AGAGCCTCAGCTATGGGACTGGTTCAATCCAAAG 630  
Db 241 GATCGCATTTGAGGAAGATGACACACAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAG 300  
QY 631 AACCCCGGATGTTTGGACAGTACCCCGTGGAAAGCGCGGATGTTGTTGGGAAGGCAC 690  
Db 301 AACCCCGGATGTTTGGACAGTACCCCGTGGAAAGCGCGGATGTTGTTGGGAAGGCAC 360  
QY 691 TATGACACAGCTCTGCTGGAAAGTACTAGCCACACAGAAACTCAGCTGTTGGGCTGACA 750  
Db 361 TATGACACAGCTCTGCTGGAAAGTACTAGCCACACAGAAACTCAGCTGTTGGGCTGACA 420  
QY 751 GTGTTTGTCTGGGAAAGTACTAGAGCATTTACTTGAAGACTTTTCTGGAGTCTGCTGAC 810  
Db 421 GTGTTTGTCTGGGAAAGTACTAGAGCATTTACTTGAAGACTTTTCTGGAGTCTGCTGAC 480  
QY 811 ATGTACTTCAATGTTGGCCATCGGGTCATATTTTACGTCTATGATAGACACACCTCCCG 870  
Db 481 ATGTACTTCAATGTTGGCCATCGGGTCATATTTTACGTCTATGATAGACACACCTCCCG 540  
QY 871 ATGCCTGTCTGCACCTGAACCTCTACATTCCTTACAAGCTTTTGGAGATCAGGTCTGAG 930  
Db 541 ATGCCTGTCTGCACCTGAACCTCTACATTCCTTACAAGCTTTTGGAGATCAGGTCTGAG 600  
QY 931 AAGAGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGCAGCATCTTGCGC 990  
Db 601 AAGAGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGCAGCATCTTGCGC 660  
QY 991 CACATCAGCAGAGGTGCGACTTCTCTTCTGATGAGCAGTGGATCAAGTCTTTTCAAGAC 1050  
Db 661 CACATCAGCAGAGGTGCGACTTCTCTTCTGATGAGCAGTGGATCAAGTCTTTTCAAGAC 720  
QY 1051 AACTTGGGCTGGAACTCTGGGCCAGCTGGTAGCAGAGCTCCAGGCTCGTGTGACAAG 1110  
Db 721 AACTTGGGCTGGAACTCTGGGCCAGCTGGTAGCAGAGCTCCAGGCTCGTGTGACAAG 780  
QY 1111 GCCAGTCCCAGAGATTTACCTATGAGAGCGGGAACCTGTGCGCGCGTACATTTCCATT 1170  
Db 781 GCCAGTCCCAGAGATTTACCTATGAGAGCGGGAACCTGTGCGCGCGTACATTTCCATT 840  
QY 1171 GGAGAGGGGATTTTACTACCACGCGGCGCATTTTGGAGGAACGCGCTACTCAGATTCT 1230  
Db 841 GGAGAGGGGATTTTACTACCACGCGGCGCATTTTGGAGGAACGCGCTACTCAGATTCT 900  
QY 1231 AACCTCACCAGGAGTGTCTTTAAGGGGATCCTCCAGGACAAGAACATGACATAGAAGCC 1290  
Db 901 AACCTCACCAGGAGTGTCTTTAAGGGGATCCTCCAGGACAAGAACATGACATAGAAGCC 960  
QY 1291 CAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCACTTAAATC 1350  
Db 961 CAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCACTTAAATC 1020  
QY 1351 CTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCGTGGCTTCTAGATATTTAAAGTGT 1410  
Db 1021 CTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCGTGGCTTCTAGATATTTAAAGTGT 1080  
QY 1411 AAGGTAGCTTTGGCAGACAAAAGAGTATAATTTGGTTAGAAATATGCTCTCA 1461  
Db 1081 AAGGTAGCTTTGGCAGACAAAAGAGTATAATTTGGTTAGAAATATGCTCTCA 1131

RESULT 9  
US-08-214-580A-2  
; Sequence 2, Application US/08214580A  
; Patent No. 5821117

GENERAL INFORMATION:  
APPLICANT: Sandrin, Mauro S.  
TITLE OF INVENTION: XENOTRANSPLANTATION  
TITLE OF INVENTION: THERAPIES  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Maurice M. Klee  
STREET: 1951 Burr Street  
CITY: Fairfield  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06430  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.4 Mb storage  
COMPUTER: Dell Dimension p166v  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Word for Windows 95  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/214,580A  
FILING DATE: 15-MARCH-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian Provisional  
APPLICATION NUMBER: Patent Application Serial No. 5821117 PL 7854  
FILING DATE: March 16, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: ALX-137  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255-1400  
TELEFAX: (203) 254-1101  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1423 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
DESCRIPTION: galactosyl transferase,  
DESCRIPTION: full coding sequence  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Sus scrofa  
US-08-214-580A-2

Query Match 48.4%; Score: 726.6; DB 1; Length 1423;  
Best Local Similarity 78.8%; Pred. No. 1.1e-199;  
Matches 893; Conservative 0; Mismatches 234; Indels 6; Gaps 2;

Qy 354 CTCAGATCTCGATCACAGGAGAAATCAATGTCAGGGAAGTAATCCCTGTTGAT 413  
Db 22 CCCAGCTTCTCCGATCAGGAGAAATCAATGTCAGGGAAGAGTGGTCTGTCAT 81

Qy 414 GCTGATTCTCTCAACCCGTTGTTGTTGGGAATATGTCACAGAAATTCACAGGT 473  
Db 82 GCTGTTCTCACTGTAATGGTTGTTGGGAATATCAACAGAAACCCAGAGT 141

Qy 474 TGGTGAGAACAGATGGCAGAGAGCTGGTGTCCCAAGCTGTTTAAATGGGACCCA 533  
Db 142 TGG---CAGCAGTGTCTCAGAGGGGCTGGTGTTCAGAGCTGGTTTAAACATGGGACTCA 198

Qy 534 CAGTTATCAAGAACACACCTAGAGGACGAGAGAAAGG---GTAGAAATGGAGATCG 590  
Db 199 CAGTTTACCAGCAAGAAAGACGCTATAGGCAACGAAAGAAACAGAAAGAGACAA 258

Qy 591 CATTGAAGAGCCTCAGCTATGGGACTGGTTCATCCAAAGAACCCCGGATGTTTTCAC 650  
Db 259 CAGAGGAGGCTCCGCTAGTGGACTGGTTAATCCTCAGAAACCCGACAGAGGTCTGAC 318

Qy 651 AGTGACCCCGTGGAGGCGCCGATTTGTGCGGAAGGCACATTATGACACAGCTGTGCTGGA 710  
Db 319 CATAACCCAGATGGAGGCTCCAGTGGTATGGAAGGCACATTACAACAGCGCTCTTGA 378

Qy 711 AAAGTACTACCCACACAGAAACCTCACTGTGGGCTGACACTGTTGCTGTGGGAACTGA 770  
Db 379 TAATATTATCCCAACAGAAATTTACCGTGGGCTTGACGGTGTTCGCTGCGGAAGATA 438

Qy 771 CATTGAGCATTACTTAGAAGACTTTCTGGAGTCTGCTGACATGTACTTTCATGGTTGGCCA 830  
Db 439 CATTGAGCATTACTTGGAGGAGTCTTAATATCTGCAATACATCTCATGTTGGCCA 498

Qy 831 TCGGGTCATATTTACGTCATGATAGACGACACCTCCCGGATGCTGTGTCACCTGAA 890  
Db 499 CAAAGTCATCTTTTACATCATGGTGGATATCTCCAGGATGCTTTGATAGAGCTGG 558

Qy 891 CCCTCTACATTCCTTACAAGCTTTTGAAGTCTTGAAGTCTGAGAGAGGTTGGCAGGATATCAG 950  
Db 559 TCCTCTCGGTTCTTTAAAGTGTGTTGAGATCAAGTCCGAGAGAGGTTGGCAGGATATCAG 618

Qy 951 CATGATGCGCATGAAGACCATTTGGGAGCAGCATCTTGCCACACATCCAGCAGGAGTCA 1010  
Db 619 CATGATGCGCATGAAGACCATCGGGAGCAGCATCTTGCCACACATCCAGCAGGAGTCA 678

Qy 1011 CTTCTCTTCTGTCATGACGTCAGTCAAGTCTTTTCAAGACAACTTCGGGGTGGAACTCT 1070  
Db 679 CTTCTCTTCTGTCATGACGTCAGTCAAGTCTTTTCAAGACAACTTCGGGGTGGAGACCT 738

Qy 1071 GGGCCAGCTGTAGCAGCAGCTCGAGGCTGTGTCACAGGCCAGTCCCGAGAGAGTTTCCAC 1130  
Db 739 GGGCCAGCTGTGTCAGCTACAGGCTGTGTCACAGGCCAGTCTTGACGAGTTCAC 798

Qy 1131 CTATGAGAGGCGGAACTGTCCGCGGCTACATTTCCATTCGGAGAGGGGGATTTTACTA 1190  
Db 799 CTAGAGAGGCGGAGGAGTCCGAGCCTACATTTCCGTTGCGCCAGGGGATTTTATTA 858

Qy 1191 CCACGCGCCCATTTTGGAGGAGCGCTACTACATCTCAACCTCACCAGGGAGTGTCT 1250  
Db 859 CCACGCGCCCATTTTGGGGAACACCCACTCAGGTCTTAAACATCACTCAGGAGTGTCT 918

Qy 1251 TAAGGGATCTCCAGGACAAAGAACATGACATAGAGCCAGTGGCATGATGAGAGCCA 1310  
Db 919 CAAGGGAATCTCCAGGACAAAGAAATGACATAGAGCCAGTGGCATGATGAGAGCCA 978

Qy 1311 CCTCAACAAATACTTCTCTTTTCAACAAACCCACTTAAATCTCTATCCAGAGTATTCGT 1370  
Db 979 TCTAAACAAGTATTTCTCTTCTCAACAAACCCACTTAAATCTTATCCCCAGAAATCTGCTG 1038

Qy 1371 GGACTATCAGATAGGCTGCTTCAGATATTTAAAGTGTCAAGGTAGCTTGGCAGACAAA 1430  
Db 1039 GGATTATCATATAGGCATGCTGTGGATATTTAGGATTTGTCAAGATAGCTTGGCAGAAA 1098

Qy 1431 AGACTATAATTTGTTAGAAATAATGTCGACTTCAAAATTTGATGGAACCTT 1483  
Db 1099 AGACTATAATTTGTTAGAAATAATGTCGACTTCAAAATTTGATGGAACCTT 1151

RESULT 10  
PCT-US95-07554-5  
Sequence 5, Application PC/TUS9507554  
GENERAL INFORMATION:  
APPLICANT: Sandrin, Mauro S.  
APPLICANT: Fodor, William L.  
APPLICANT: Rother, Russell P.  
APPLICANT: Squinto, Stephen P.  
APPLICANT: McKenzie, Ian F. C.  
TITLE OF INVENTION: Methods for Reducing  
Hyperacute Rejection of Xenografts  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Maurice M. Klee  
STREET: 1951 Burr Street  
CITY: Fairfield

STATE: Connecticut  
COUNTRY: USA  
ZIP: 06430  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 5 inch, 750 kb storage  
COMPUTER: Dell 486/50  
OPERATING SYSTEM: DOS 6.2  
SOFTWARE: WordPerfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07554  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/260,201  
FILING DATE: June 15, 1994  
CLASSIFICATION:  
APPLICATION NUMBER: 08/278,282  
FILING DATE: July 21, 1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: ALX-144.1PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255-1400  
FAX: (203) 254-1101  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1423 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
DESCRIPTION: galactosyl transferase,  
DESCRIPTION: full coding sequence  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: Sus scrofa  
PCT-US95-07554-5

Query Match 48.4%; Score 726.6; DB 5; Length 1423;  
Best Local Similarity 78.8%; Pred. No. 1.1e-199;  
Matches 893; Conservative 0; Mismatches 234; Indels 6; Gaps 2;

Qy 354 CTCAGATCTGGATCACAGGAAATAATCAATGTCAAGGAAAGTAATCTCTGTGAT 413  
Db 22 CCCAGCTCTGCCGATCAGGAGAAATAATGAATGTCAAAGGAAGAGTGTCTGTCAAT 81

Qy 414 GCTGATTGTCTCAACCGTGTGTGTGTTGGGAATATGTCAACAGAATTCAGAGGT 473  
Db 82 GCTGTTGTCTCACTGTAATGTGTGTGTTGGGATATACATCAACAGAAACCCAGAGT 141

Qy 474 TGGTGAACAGATGGCAGAGGACTGGTGTCCCAAGCTGGTTTAAAAATGGGACCCA 533  
Db 142 TGG---CAGCAGTCTCAGAGGGGCTGGTGTTCGAGCTGGTTTAAACAATGGGACTCA 198

Qy 534 CAGTTATCAAGAACACACGCTAGAAGCAGGAGAGAAAGS---GTAGAAATGGAGATCG 590  
Db 199 CAGTTACACGAGAAAGACGCTATAGGACGAAAGAAAGAAAGAAAGAAAGACAA 258

Qy 591 CATTGAAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAGAACCGCCGGATGTTTGAC 650  
Db 259 CAGAGGAGAGCTCCGCTAGTGGACTGTTAATCCTTGAGAAACGCCAGAGGTCTGTAC 318

Qy 651 AGTACCCCGTGAAGCGCGATGTGTGGGAAGGCATTTAGACACAGCTCTGTGGA 710  
Db 319 CATAAACAGATGGAAGGCTCAGTGTGTGGGAAGGCATTTACAAACAGAGCCGCTTGA 378

Qy 711 AAGTACTACCCACACAGAACTCACTGTGGGCTCAGCTGTTCTGTGGGAAGTA 770  
Db 379 TAATTATTATGCCAACAGAAATACCGTGGGCTTACCGGTTTCTGTGCGAAGATA 438

Qy 771 CATTGAGCATTTACTTAGAAGACTTTCTGGAGTCTGCTGACATCTACTTCTATGCTGGCCA 830  
Db 439 CATTGAGCATTTACTTGAGGAGTCTTAATATCTGCAATACATCTTCTATGCTGGCCA 498

Qy 831 TCGGTCATATTTTACGTATGATAGACACACCTCCCGGATCGCTGCTGCACTGAA 890  
Db 499 CAAAGTCATCTTTTACATCATGCTGATATCTCCAGGATGCTTTGATAGAGCTGG 558

Qy 891 CCCTCTACATCTCTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGTGGCAGATATCAG 950  
Db 559 TCCTCTGCTCTCTTTAAAGTGTTCGAGATCAAGTCCGAGAAGAGTGGCAACATCAG 618

Qy 951 CATGATCGCATGAAGACCATTTGGGAGCAGATCTTGGCCACATCCAGCAGAGTCCA 1010  
Db 619 CATGATCGCATGAAGACCATCGGGAGCAGATCTTGGCCACATCCAGCAGAGTCCA 678

Qy 1011 CTTCCTCTCTGATGAGCAGTGTCAAGTCTTTCAAGACAACCTTCGGGCTGGAACTCT 1070  
Db 679 CTTCCTCTCTGATGAGCAGTGTCAAGTCTTTCAAGACAACCTTCGGGCTGGAGCTT 738

Qy 1071 GGCCAGCTGTAGCAGCAGCTCCAGGCTGTGTGTCAAGGCCAGTCCCGAGAGTTCAC 1130  
Db 739 GGCCAGCTGT 798

Qy 1131 CTATGAGAGCGGGAACATGTGGCCGCGTACATTTCCATTCGAGAGGGGATTTTACTA 1190  
Db 799 CTAGGAGCGGGAAGGAGTCCGAGCTACATTTCCGCTTTGGCAGGGGATTTTATTA 858

Qy 1191 CCAGCGGCTCTTTTGGAGGAGCGCTACTACATTTCTCAACCTCAGGAGGAGTCTT 1250  
Db 859 CCAGCGGCTCTTTTGGGGAACACCCACTCAGGTCTTAAACATCACTCAGGAGTCTT 918

Qy 1251 TAAGGGATCTCCAGGACAAAGAAATGACATAGAGCCAGTGGATGATGAGAGCCA 1310  
Db 919 CAAGGGAATCTCCAGGACAAAGAAATGACATAGAGCCAGTGGATGATGAGAGCCA 978

Qy 1311 CCTCAACAAATCTCTCTTTTCAACAAACCCACTTAAATCTATCTCCAGAGTATTCGTG 1370  
Db 979 TCTAAACAAGTATTTCTCTTCAACAAACCCACTTAAATCTTATCCCAAGATCTGCTG 1038

Qy 1371 GGACTATCAGATAGGCTGCTTCAGATATTTAAAGTGTCAAGTGTCTGGCAGACAAA 1430  
Db 1039 GGATATATATAGGCTGCTGCTGGATATAGGATTTGTCAAGATCTTGGCAGACAAA 1098

Qy 1431 AGAGTAAATTTGGTTAGAATAATGTCTGACTTCAAAATTTGTGATGAAACTT 1483  
Db 1099 AGAGTAAATTTGGTTAGAATAATGTCTGACTTAAATTTGTGCTCCAGCAGTCTT 1151

RESULT 11  
US-08-704-548-1  
Sequence 1 Application US/08704548  
Patent No. 5879675  
GENERAL INFORMATION:  
APPLICANT: GALILI, URI  
APPLICANT: REPIK, PATRICIA M.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR VACCINES  
TITLE OF INVENTION: COMPRISING ALPHA-GALACTOSYL EPITOPES  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEIDEL, GONDA, LAVORNA & MONACO, P.C.  
STREET: Suite 1800, Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/704,548  
FILING DATE: 11-SEP-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Moraco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 8760-2 CII  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1140 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 10..1140  
US-08-704-548-1

Query Match 46.98; Score 702.8; DB 2; Length 1140;  
Best Local Similarity 78.28; Pred. No. 7.4e-193;  
Matches 892; Conservative 0; Mismatches 197; Indels 51; Gaps 2;

Qy 373 GAGAAATATGAATGTCACAGGAAAGTAACTCTGTTGATGCTGATGCTCAACCGTG 432  
Db 1 GAGAAATATGAATGTCACAGGAAAGTAACTCTGTCGATGCTGCTCAACTGTG 60

Qy 433 GTTCGCTGTTTTGGGAATATGTCAACAGA----- 462  
Db 61 ATTCTGTGTTTTGGGAATATATCAACAGCCAGAGGCTCTTCTTCTGGATATATCAC 120

Qy 463 -----ATTCAGAGGTGGTGACACAGATGGCAGAGGACTGGTGTCCCAAGCTGG 516  
Db 121 TCAAGAACCCAGAGGTGGTGACAGCAGTCTCAGAGGACTGGTGTTCGGCTGG 180

Qy 517 TTTAAATGGGACCCACAGTTATCAAGAGACACAACTAGAGGACGAGAGAGAGAAAGGT 576  
Db 181 TTTAAATGGGACCCACAGTTATCAAGAGAGACACAACTAGAGGACGAGAGAGAGAAAGGA 240

Qy 577 AG-----AAATGGAGATGCGATTGAAGAGCCTCAGCTATGGACTGGTTC 621  
Db 241 AGAGAGGAGGAACAAAAGAGAGATGACACAAACAGAGCTTCGGCTATGGGACTGGTTT 300

Qy 622 AATCAAGAACCCGCGGATGTTTTCACAGTGACCCGCTGGAGGCGCGGATGTTGG 681  
Db 301 AATCAAGAACCCGCGGATGTTTTCACAGTGACCCGCTGGAGGCGCGGATGTTGG 360

Qy 682 GAAGGCACTTATGACACAGCTCTGCTGAAAGTACTACGACACACAGAACTCACTGTG 741  
Db 361 GAAGGCACTTATGACACAGCTCTGCTGAAAGTACTACGACACACAGAACTCACTGTG 420

Qy 742 GGGCTGACAGTGTGTTGCTGGGAAAGTACATGAGGACTTACTTGAAGACTTCTTGAG 801  
Db 421 GGGTTGACGTTTGTGCTATTTGGAAGATATATGAGCATTTACTTGGAGGACTTCGTAACA 480

Qy 802 TCTGCTGACATGACTTTCATGTTGGCCATCGGCTCATATTTACGTCATGATAGACGAC 861  
Db 481 TCTGCTGACATGACTTTCATGTTGGCCATCGGCTCATATTTACGTCATGATAGACGAC 540

Qy 862 ACCTCCGGATGCTGTGCTGACCTGAACCTCTACATCTCTTACAACTCTTTGAGATC 921  
Db 541 GTCTCCAGGCGCGTTTATAGAGCTGGTCTCTGCTTCTTCAAGTGTTCGAGTC 600

Qy 922 AGGCTGAGAGAGGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGCAC 981  
Db 601 AAGCCAGAGAGAGGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGCAC 660

Qy 982 ATCTGCGCCACATCCAGCAGGAGTGGCTCTCTCTGTCATGGAGCTGGATCAAGTC 1041  
Db 1041 ATCTGCGCCACATCCAGCAGGAGTGGCTCTCTCTGTCATGGAGCTGGATCAAGTC 1041

Db 661 ATCTTGGCCACATCCAACACAGAGTTTACCTTCTTCTGCATGGATGTGGACAGGTC 720  
Qy 1042 TTTCAGACAACTTTCGGGTGGAACCTCTGGCCAGCTGGTAGCACAGCTCCAGGCTGG 1101  
Db 721 TTCCAAGACCAATTTTGGGTAGACACCTGGCCAGCTGGTGGCTCAGCTACAGGCTGG 780  
Qy 1102 TGGTACAAAGCCAGTCCCGAGAAAGTTTCACTTATGAGAGGCGGGAACCTGTCGGCGCGTAC 1161  
Db 781 TGGTACAAAGCCAGATCCTGTGACTTTTACCTTATGAGAGGCGGAAAGAGTCCGCGACATAT 840  
Qy 1162 ATTCCATTCGAGAGGCGGATTTTACTTACACACGCGGCGCATTTTGGAGGAACGCTACT 1221  
Db 841 ATTCCATTTGGCCAGGCGGATTTTATATACCATGACGCCATTTTGGAGGAACCCGAT 900  
Qy 1222 CACATTCCTCAACCTCACAGGAGTGTCTTAAAGGGATCTCCAGGACAAACAAATGAC 1281  
Db 901 CAGTTTCTCAACATCACCCAGGAGTGTCTTAAAGGAATCTCTCCGACAAAGAAATGAC 960  
Qy 1282 ATAGAAGCCCAAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTCTTCAACAAACCC 1341  
Db 961 ATAGAAGCCGAGTGGCATGATGAGAGCCACCTTAAAGGACCTTAAACAAAGTATTTCTTCTCAACAAACCC 1020  
Qy 1342 ACTAAATCTCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCTTTCAGATATT 1401  
Db 1021 TCTAAATCTTATCTCCAGAACTACTGCTGGGATTTATCATATAGGCTGCTTTCAGATATT 1080  
Qy 1402 AAAAGTGTCAAGTAGCTTGGCAGACAAAGAGATATAATTTGGTTAGAAATATGTCTGA 1461  
Db 1081 AAAACTGTCAAGCTATCATGCGCAACAAAGAGATATAATTTGGTTAGAAAGATGTCTGA 1140

RESULT 12  
US-08-621-700-1  
: Sequence 1, Application US/08621700  
: Patent No. 6153428  
: GENERAL INFORMATION:  
: APPLICANT: Gustafsson, Kenneth T.  
: Sacha, David. H.  
: TITLE OF INVENTION: '(1,3) GALACTOSYLTRANSFERASE NEGATIVE SWINE  
: NUMBER OF SEQUENCES: 24  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Carella, Bryne, Bain, Gilfillan, Cecchi, Stewart & Olstein  
: STREET: 6 Becker Farm Road  
: CITY: Roseland  
: STATE: New Jersey  
: COUNTRY: USA  
: ZIP: 07068  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: 3.5 inch diskette  
: COMPUTER: IBM PS/2  
: OPERATING SYSTEM: MS-DOS  
: SOFTWARE: ASCII  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/621,700  
: FILING DATE: 26-Mar-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 08/228,933  
: FILING DATE: <Unknown>  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Charles J. Herron  
: REGISTRATION NUMBER: 28,019  
: REFERENCE/DOCKET NUMBER: 61750-104  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1269 bases  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: HYPOTHETICAL: NO  
: ANTI-SENSE: NO  
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-08-621-700-1

Query Match 45.6%; Score 684.2; DB 3; Length 1269;  
Best Local Similarity 76.6%; Pred. No. 1.8e-187;  
Matches 880; Conservative 0; Mismatches 233; Indels 36; Gaps 2;

QY 371 AGGAGAAATAATGATGTCAGGGAAGAGTAATCTGTGTGATGCTGATGCTCAACCG 430  
DB 5 AGGAGAAATAATGATGTCAGGGAAGAGTGTGTTCTGCAATGCTGTGCTCAACTG 64  
QY 431 TGGTGTGCTGTTTGGGAATATGTCACAGAAATTCAGAGGTT----- 474  
DB 65 TAATGGTTGTGTTTGGGAATATGTCACAGAAATTCAGAGGTT----- 474  
QY 475 -----GGTGAAGACAGATGGCAGAGGAGTGTGTTTCCCAAGCTGGT 517  
DB 125 AGTCAAAAACCCAGAAAGTTGGCAGCAGTGCTCAGAGGCGCTGTTTCCAGCTGGT 184  
QY 518 TTAATAATGGGACCCACAGTTATCAAGAGACACAACTAGAGGACGAGAGAAAGG--- 574  
DB 185 TTAACAAATGGGACTACAGTTACGAGAGAGAGCTTCCGCTAGTGGCTTAACTCTGAAAC 244  
QY 575 GTAGAAATGGAGATCGCATTTGAAGACGCTCAGCTATGGGACTGTTCAATCCAAAGAAC 634  
DB 245 AAGAAAGAGACACAGAGGAGAGCTTCCGCTAGTGGCTTAACTCTGAAAC 304  
QY 635 GCCCGGATGTTTTCAGAGTGACCCGCTGGAAGCGCCGATTTGTGGGAGGACACTTATG 694  
DB 305 GCCCAGAGGTCGTGACCAATCAACAGATGGAAGGCTCCAGTGATGGGAGGACACTTACA 364  
QY 695 ACACAGCTCTGCGGAAAGTACTAGCCACACAGAAACTCACTGTGGGCTGACAGTGT 754  
DB 365 ACAGAGCCGCTCTAGATAATATTTATGCCAAACAGAAATATACCGTGGGCTTGACGGTTT 424  
QY 755 TTCTCTGTGGGAAGTACATTTGAGCATTACTTTAGAAGCTTTCTGGAGTCTGCTGACATGT 814  
DB 425 TTCTGTGCGGAAGATACATTTGAGCATTACTTTGAGGAGTCTTAAATATCTGCAAAATACAT 484  
QY 815 ACTTCATGTTGGCCATCGGGTCATATTTTACGTCATGATAGACACACTCCCGGATGC 874  
DB 485 ACTTCATGTTGGCCACAAAGTCACTTTTACATCATGTTGGATGATATCTCCAGGATGC 544  
QY 875 CTCTCTGTGCACCTGAACCCCTCTACATCTCTTACAAGCTTTTGAGATCAGGTCTGAGAAGA 934  
DB 545 CTTTGTATGAGAGCTGGTCTCTCGTCTCTTAAAGTGTTTGAGATCAAGTCCGAGAAGA 604  
QY 935 GTGGCAGGATATCAGCATGATCGCATGAAGACCATTTGGGAGACATCTCGGCCACAC 994  
DB 605 GTGGCAAGACATCAGCATGATCGCATGAAGACCATCGGGAGACATCTCGGCCACAC 664  
QY 995 TCCAGCACGAGTTCGACTTCTCTGTCATGGACGTGGATCAAGTCTTTTCAAGACAAC 1054  
DB 665 TCCAGCACGAGTTCGACTTCTCTGTCATGGACGTGGATCAAGTCTTTTCAAAACAAC 724  
QY 1055 TCCGGGTGGAAACTCTGGGCCACTGGTACGACAGCTCCAGGCTGGTCAAGGCCCA 1114  
DB 725 TTGGGTGGAGACCTTGGGCCACTGGTGGCTCAGTACAGGCTGGTGTACAGGCCAC 784  
QY 1115 GTCCCGAAGATTTCACTATGAGAGCGGGAAGTGTTCGGCCGGTACATTTCCATTCGGAG 1174  
DB 785 ATCTGACGAGTTTCACTAGAGAGCGGGAAGAGTCCCGAGGCTACATTTCCGTTTGGCC 844  
QY 1175 AGGGGATTTTTTACTACGAGCGGCCATTTTGGAGGAAGCGCTACTTCATCTTCAAC 1234  
DB 845 AGGGGATTTTTTATACGAGCGGCCATTTTGGGGAACACCCACTCAGGTTCTAAACA 904  
QY 1235 TCACCGAGGAGTCTTTAAGGGATCTCTCAGGACAGAAACATGACATAGAAGCCAGT 1294  
DB 905 TCACTCAGGAGTCTTTCAAGGGAAATCTCTCAGGACAGGAAATGACATAGAAGCCGAG 964  
QY 1295 GGCATGATGAGCCACCTTCAACAAATACTTCTCTTTTCAACAAACCCACTTAAATCCTAT 1354  
DB 965 GGCATGATGAAAGCCATCTTAAACAAGTATTTCTCTTCTCAACAAACCCACTTAAATCTTAT 1024

QY 1355 CTCCAGAGATATTCTCTGGGACTATCAGATAGGCTGCTTTCAGATATAAAGTGTCAAGG 1414  
DB 1025 CCCCAAGATAGCTGGGATTTATCATATAGGCATGCTCTGGGATATTAGGATTTGTCAAGA 1084  
QY 1415 TAGCTTGGCAGACAAAGAGTATTAATTTGGTTAGAATAATGCTCTGACTTCAAAATTGTCA 1474  
DB 1085 TAGCTTGGCAGAAAAAGAGTATTAATTTGGTTAGAATAATCAATCTGACTTAAATTTGTGC 1144  
QY 1475 TGGAAACTT 1483  
DB 1145 CAGCAGTTT 1153

RESULT 13  
PCT-US95-03940-1  
; Sequence 1, Application PC/TUS9503940  
; GENERAL INFORMATION:  
; APPLICANT: BIOTRANSPLANT, INC.  
; APPLICANT: THE GENERAL HOSPITAL CORPORATION  
; APPLICANT: THE UNIVERSITY OF LONDON  
; APPLICANT: BAETSCHER, Manfred W.  
; APPLICANT: GUSTAFSSON, Kenth T.  
; APPLICANT: SACHS, David H.  
; TITLE OF INVENTION: (1,3) GALACTOSYLTRANSFERASE NEGATIVE SWINE  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,  
; ADDRESSEE: Stewart & Olstein  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/03940  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/228,933  
; FILING DATE: April 13, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Charles J. Herron  
; REGISTRATION NUMBER: 28,019  
; REFERENCE/DOCKET NUMBER: 61750-CIP  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1269 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
PCT-US95-03940-1

Query Match 45.6%; Score 684.2; DB 5; Length 1269;  
Best Local Similarity 76.6%; Pred. No. 1.8e-187;  
Matches 880; Conservative 0; Mismatches 233; Indels 36; Gaps 2;

QY 371 AGGAGAAATAATGATGTCAGGGAAGAGTAATCTGTGTGATGCTGATGCTCAACCG 430  
DB 5 AGGAGAAATAATGATGTCAGGGAAGAGTGTGTTCTGCAATGCTGTGCTCAACTG 64  
QY 431 TGGTGTGCTGTTTGGGAATATGTCACAGAAATTCAGAGGTT----- 474  
DB 65 TAATGGTTGTGTTTGGGAATATGTCACAGAAATTCAGAGGTT----- 474  
QY 475 -----GGTGAAGACAGATGGCAGAGGAGTGTGTTTCCCAAGCTGGT 517

Db 125 AGTCAAAAAAAGTTGGCAGCAGTGTCTCAGAGGGGCTGGTGGTTCCTCCAGCTGGT 184  
Qy 518 TTAATAATGGGACCCAGTATCAAGAACACAGCTAGAGGAGCGGAGAGAAAGG--- 574  
Db 185 TTAAATATGGGACTCAGCTATCCACGACAGAGAGAGCTATAGGACAGAAAGGAAC 244  
Qy 575 GTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAGAAC 634  
Db 245 AAAGAAAGAGAACACAGAGGAGAGCTCCGCTAGTGGACTGGTTTAACTCCTGAGAAC 304  
Qy 635 GCCGGATGTTTGGACAGTGAACCCGTTGGAAGGCGCGATTTGTTGGGAGGACCTTATG 694  
Db 305 GCCAGAGGTGCTGACCATCAACAGATGGAAGGCTCCAGTGGTATGGGAAGGCACTTACA 364  
Qy 695 ACACAGCTCTGCTGGAAGAGTACTACCCACACAGAAACTCACTGTGGGCTGACAGCTGT 754  
Db 365 ACAGAGCGGTCTTAGATTAATTTATGCCAACAAGAAATATACGTGGGCTTGACGGTTT 424  
Qy 755 TTGCTGTGGGAAAGTACATTTGAGCATTTACTTAGAAGACTTTCTGGAGTCTGCTGACATGT 814  
Db 425 TTGCTGTGCGAAGATACATTTGAGCATTTACTTTGGAGGAGTTCTTAATATCTGCAAAATACAT 484  
Qy 815 ACTTCATGTTGGCCATCGGGGTGATATTTTACGTCTGATAGAGGACACCTCCCGGATGC 874  
Db 485 ACTTCATGTTGGCCACAAAGTCTATCTTTTACATCATGTTGGTGGATGATATCTCCAGGATGC 544  
Qy 875 CTGCTGTGACCTGAAACCTCTACATTTCTTACAACTTTTGGAGTCACTGAGCTCTGAGAGA 934  
Db 545 CTTGTGATAGAGTGGGTCTCTGGGTCTCTTTAAAGTTTGTGGATCAAGTCCGAGAGA 604  
Qy 935 GGTGGCAGGATATCAGCATATGCGATGATGAGGACCAATTTGGGAGCAGCATCTCGGCCACA 994  
Db 605 GGTGGCAAGACATCAGCATATGCGATGATGAGGACCAATTCGCGGAGCAGCATCTCGGCCACA 664  
Qy 995 TCNAGCAGAGGTGCGACTTCTCTTCTGCTGATGAGCTGGATCAAGTCTTTCAAGACAACT 1054  
Db 665 TCCAGCAGAGGTGGAGTCTCTCTCTGCTGATGAGCTGGATCAAGTCTTTCAAAACAACT 724  
Qy 1055 TCGGGTGGAAACTCTCGGCGAGTCTGAGCAGCTCCAGGCTGGTGGTACAAAGCCA 1114  
Db 725 TTGGGTGGAGACCTCGGCGAGTCTGGTGGCTCAGCTACAGGCTGGTGGTACAAAGCAC 784  
Qy 1115 GTCCCGAAGATTACCTATGAGAGCGGGAATGTGCGCGCGGTACATTTCCATTCGGAG 1174  
Db 785 ATCTGACGAGTTTCACTACGAGCGGGAAGGAGTCCGCGAGCTACATTTCCGTTTGGCC 844  
Qy 1175 AGGGGATTTTACTACCGCGGCAATTTTGGAGGACGCTTACTCAGATTTCTCAACC 1234  
Db 845 AGGGGATTTTATACCGCGAGCAATTTTGGGGAACACCCACTCAGGTTCTAAACA 904  
Qy 1235 TCACGAGGAGTGTCTTTAAGGGATCTCCAGGACAAAGAAACATGACATAGAAGCCAGT 1294  
Db 905 TCACTCAGGAGTGTCTTCAAGGGAATCTCCAGGACAAAGAAATGACATAGAAGCCAGT 964  
Qy 1295 GGCATGATGAGAGCCACTCAACAAATATCTTCTTTTCAACAAACCCACTAAAATCCTAT 1354  
Db 965 GGCATGATGAAAGCCATCAACAAAGTATTTTCTCTCAACAAACCCACTAAAATCCTAT 1024  
Qy 1355 CTCCAGAGTATTTGCTGGACTATCAGATAGGCTGCTTCCAGATTTAAAGTGTCAAGG 1414  
Db 1025 CCCAGAAATATCTGCTGGATATCATATAGGCAATGTCTGGATATTTAGGATTTGTCAGA 1084  
Qy 1415 TAGCTTGGCAGACAAAAGAGTATAATTTGGTTTGAATAATATGCTGACTTCAAAATCTGA 1474  
Db 1085 TAGCTTGGCAGAAAAGAGTATAATTTGGTTAGAAATAACATCTGACTTAAATCTGC 1144  
Qy 1475 TGGAACTT 1483  
Db 1145 CAGCAGTTT 1153

RESULT 14

US-08-378-617A-7  
; Sequence 7, Application US/08378617A  
; Patent No. 5849991  
; GENERAL INFORMATION:  
; APPLICANT: d'Apice, Anthony J.F.  
; APPLICANT: Pearse, Martin J.  
; APPLICANT: Robins, Allan J.  
; APPLICANT: Crawford, Robert J.  
; APPLICANT: Rachen, Peter D.  
; TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF  
; TITLE OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 120 South Sixth Street, Suite 2500  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/378,617A  
; FILING DATE: 26-JAN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 06868/005001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (612) 335-5070  
; TELEFAX: (612) 288-9696  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1412 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-378-617A-7

Query Match 45.6%; Score 684.2; DB 2; Length 1412;  
Best Local Similarity 76.6%; Pred. No. 1.9e-187;  
Matches 880; Conservative 0; Mismatches 233; Indels 36; Gaps 2;

Qy 371 AGGAGAAAATAATGAATGTCAAGGGAAAAGTAATCTCTGTGTGATGCTGATTTGCTCAACCG 430  
Db 80 AGGAGAAAATAATGAATGTCAAGGGAAAAGTGGTCTCTCAATGCTGTTGCTCAACTG 139  
Qy 431 TGGTGTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTT----- 474  
Db 140 TAATGGTGTGTTTGGGAATATCATCAACAGCCCGCAGAGGTTCTTTCTTCTGGATATACC 199  
Qy 475 -----GGTGAGAACAGATGGCAGAGGACTGCTGTCTCCCAAGCTGGT 517  
Db 200 AGTCAAAAACCCAGAGTTGGCAGCAGTCTCAGAGGGGCTGGTGTTCGAGCTGGT 259  
Qy 518 TTAATAATGGAGCCAGCTTATCAAGAAGACACACGTAGAGGACGAGAGAGAAAGG--- 574  
Db 260 TTAACAAATGGAGCTCAGCTTACCACCAAGAAAGACGCTATAGGCAAGCAAGGAAC 319  
Qy 575 GTAGAAATGAGATCGCATTTGAAGAGCCTCAGCTATCGGACTGTTCAATCCAAAGAAC 634  
Db 320 AAAGAAAGAGAACACAGAGAGGAGCTTCGCTAGTGGACTGGTTTAAATCCTGAGAAC 379  
Qy 635 GCCCGATGTTTTCAGAGTGACCCCGGTGGAAGGCGCGATTTGCTGGGAGGCACTTATG 694  
Db 380 GCCCAGAGGTCTGACCATTAACAGATGGAAGGCTCCAGTGGTATGGAGGCACTTACA 439







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 17:56:14 ; Search time 252.09 Seconds  
(without alignments)  
10216.080 Million cell updates/sec

Title: US-09-863-475a-3  
Perfect score: 1500  
Sequence: 1 CTTTCCCTGTAGACTTTC.....CTTGACACTATTCTTAACCA 1500

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues  
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				N_Geneseq_032802.*	
1:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*	1	1500	100.0	1500
2:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*	2	1500	100.0	1500
3:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*	3	1500	100.0	1500
4:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*	4	1500	100.0	1500
5:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*	5	1500	100.0	1500
6:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*	6	1500	100.0	1500
7:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*	7	1500	100.0	1500
8:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*	8	1500	100.0	1500
9:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*	9	1500	100.0	1500
10:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*	10	1500	100.0	1500
11:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*	11	1500	100.0	1500
12:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*	12	1500	100.0	1500
13:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*	13	1500	100.0	1500
14:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*	14	1500	100.0	1500
15:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*	15	1500	100.0	1500
16:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*	16	1500	100.0	1500
17:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*	17	1500	100.0	1500
18:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*	18	1500	100.0	1500
19:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*	19	1500	100.0	1500
20:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*	20	1500	100.0	1500
21:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*	21	1500	100.0	1500
22:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*	22	1500	100.0	1500
23:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*	23	1500	100.0	1500
24:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*	24	1500	100.0	1500

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1500	100.0	1500	12	AAQ13331
2	1500	100.0	1500	15	AAQ56907
3	1500	100.0	1500	18	AAT61676
4	1500	100.0	1500	24	AAD26484
5	1434.8	95.7	3537	22	AAS05315
6	1030.8	68.7	1116	16	AAT04523
7	726.6	48.4	1423	15	AAQ74712
8	726.6	48.4	1423	17	AAT12242
9	725.2	48.3	3240	22	AAS05314

10	706	47.1	1092	19	AAV49454	Porcine alpha-1,3-
11	693.8	46.3	1131	16	AAT04522	Marmoset alpha-1,3
12	693.8	46.3	1131	24	AAD26482	Marmoset alpha (1,
13	684.2	45.6	1269	24	AAD26483	Pig alpha (1, 3) g
14	681	45.4	1269	16	AAT02892	Porcine alpha (1,3
15	677.8	45.2	1410	16	AAQ93077	Pig alpha-1,3-gala
16	664.6	44.3	1128	19	AAD49453	Porcine alpha-1,3-
17	645.6	43.0	1496	24	AAD26480	Sheep alpha (1, 3)
18	642.4	42.8	1617	24	AAD26481	Bovine alpha (1, 3
19	632.8	42.2	1065	19	AAV49455	Porcine alpha-1,3-
20	617.2	41.1	1029	19	AAV49456	Porcine alpha-1,3-
21	583.2	38.9	1353	15	AAQ74711	Galactosyl transfe
22	561.8	37.5	1885	24	AAD26485	Human alpha (1, 3)
23	518.2	34.5	2964	22	AAS05303	Pig alpha-1,3 gala
24	516.2	34.4	3322	22	AAS05322	Human alpha-1,3 ga
25	511.2	34.1	3745	22	AAS05323	Human alpha-1,3 ga
26	479	31.9	3135	22	AAS05316	Human alpha-1,3 ga
27	479	31.9	3538	22	AAS05317	Human alpha-1,3 ga
28	477.4	31.8	2303	22	AAS05332	Rhesus monkey alph
29	476.6	31.8	2848	22	AAS05331	Human alpha-1,3 ga
30	475.6	31.7	2630	22	AAS05333	Rhesus monkey alph
31	241.6	16.1	826	20	AAV65174	Human Huma exon 6
32	241.6	16.1	1593	22	AAQ09959	Human drug metabol
33	240	16.0	1062	12	AAQ11131	Histo-blood gp. A
34	240	16.0	1062	15	AAQ68806	Human A transferas
35	233.6	15.6	826	20	AAV65175	Baboon A allele DN
36	232	15.5	826	20	AAV65176	Baboon A allele DN
37	232	15.5	826	20	AAV65177	Baboon B allele DN
38	211	14.1	1684	22	AAI58407	Human polynucleoti
39	188.6	12.6	585	15	AAQ68826	Partial sequence o
40	184.4	12.3	900	22	AAS05308	Mouse alpha-1,3 ga
41	184.2	12.3	1926	22	AAI60193	Human polynucleoti
42	182.6	12.2	1926	22	AAK94196	Human full-length
43	164.4	11.0	2718	22	AAK94349	Human full-length
44	141.8	9.5	608	22	AAS05313	Mouse alpha-1,3 ga
45	110	7.3	1207	22	AAS05307	Mouse alpha-1,3 ga

ALIGNMENTS

RESULT 1	AAQ13331	standard; cDNA; 1500 BP.
ID	AAQ13331	standard; cDNA; 1500 BP.
XX	AAQ13331;	
AC	AAQ13331;	
DT	07-NOV-1991	(first entry)
XX	UDP-Gal:beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-	
DE	galactosyltransferase gene.	
DE	Glycosyltransferase; ss.	
XX	Mus musculus.	
OS	Mus musculus.	
XX	Key	Location/Qualifiers
FH	CDS	277..1185
FT		/*tag= a
XX	WO9112340-A.	
PN	22-AUG-1991.	
XX	14-FEB-1991;	91WO-US00899.
XX	12-DEC-1990;	90US-0627621.
PR	14-FEB-1990;	90US-0479858.
PR	14-FEB-1990;	90US-0480133.
XX	(UNMI ) UNIV OF MICHIGAN.	
PA		
XX	Lowe JB;	
PI		





XX	Alpha(1,3)-galactosyltransferase; glycosylation; oligosaccharide;
KW	ss.
XX	
OS	Mus sp.
XX	
FH	Location/Qualifiers
FT	277..1461
FT	/*tag= a
XX	
PN	WO9709421-A1.
XX	
PD	13-MAR-1997.
XX	
PF	06-SEP-1996; 96WO-US13816.
XX	
PR	08-SEP-1995; 95US-0525058.
XX	
PA	(UNMI ) UNIV MICHIGAN.
XX	
PI	Legault DJ, Lowe JB;
XX	
DR	WPI: 1997-192897/17.
DR	P-PSDB; AAW13639.
XX	
PT	New recombinant fucosyltransferase proteins - useful for modifying
PT	cell surface oligosaccharide structures
XX	
PS	Example 2; Page 271-272; 329pp; English.
XX	
CC	A cDNA clone (AAT61676) codes for an enzyme (AAW13639) that functions
CC	as a UDP-Gal:beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-
CC	galactosyltransferase. It was obt'd. by transfecting COS-1 cells
CC	with cDNA derived from mouse F9 teratocarcinoma cells, and
CC	screening the transfected cells for surface-localised
CC	Gal(alpha1-3)Gal linkages. The cDNA can be used to construct
CC	animal cell lines with specific capabilities with respect to post-
CC	translational modification of the oligosaccharides of expressed
CC	proteins or lipids, or to produce recombinant enzyme for use in
CC	oligosaccharide prodn.
XX	
SQ	Sequence 1500 BP; 406 A; 349 C; 374 G; 371 T; 0 other;
Query Match 100.0%; Score 1500; DB 18; Length 1500;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 CCTTCCCTTGTACACTCTTCTTGGATGAGAACTACCGATTCTGCTGAAGACCTCGCGCT 60
DB	1 ccttcccttgtagactctcttggaatgagaagtacccgattctgctgaagaccctgcgct 60
QY	61 CTCAGGCTCTGGAGTTGGAACTCTGACCTTCTTCCCTGCTGAGCCCTGCCTCCTT 120
DB	61 ctcaggctctggaggttggaaacctgtaccctctcttctctctgctgagccctgcctctt 120
QY	121 AGCCAGCCAGAGCTCCACAGAACTCGGTTCGTTGCTTGTGTTGGAGGGAACACAG 180
DB	121 aggcaggccagagctgcagagaactcggttgcttctgttctgttggagggaacacag 180
QY	181 CTGACGATGAGGTGACTTTGAACCTCAAGAGATCTGCTTACCCAGCTCTCCTGGAATTAA 240
DB	181 ctgacgatgaggctgactttgaaactcaagagatctgtaccaccagctctctcgtgaaattaa 240
QY	241 AGCCCTGTACTACATTTGCTGGACCTCAAGATTTTCATGATCACTATGCTPTCAAGATCTC 300
DB	241 agccctgtactacatttgcctggacctaaagattttcatgatcactatgcttcaagatctc 300
QY	301 CATGTCACAAGATCTCCATGTCCTCAAGATCCAAAGTCAAGAAACAAGTCTTCCATCCTCAAGA 360
DB	301 catgtcaacaagatctccatgtcaagatcccaagtcagaacaagctcttccatcctcaaga 360
QY	361 TCTGGATCACAGGAGAAATAATGAATGTCAAGGAAAGTAATCCTGTGTGATGCTGATT 420

DB	361 tctggatcacaggagaaaaataatgaatgtcaagggaagaagtaactcctgttgtatgctgatt 420
QY	421 GTCTCAACCGTGTGTCGTTTGGGAATATGCTCAACAGAAATTCACAGAGTTGGTGAG 480
DB	421 gtctcaaccgtggtcgtcgtcgttctgggaatactgcacagaaatcccaagggttcggcgag 480
QY	481 AACAGATGGCAGAGGACTGCTGCCAAGCTGTTTAAATAATGGACCCACAGTTAT 540
DB	481 aacagatggcagaaggactggttcccaagctggttttaaaatggagaccacagttat 540
QY	541 CAAGAGACAACGTAGAGACGGAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAG 600
DB	541 caagaagacaacgtagaaggacgagagaaaggtagaaatggagatcgcatttgaagag 600
QY	601 CTTACGCTATGGGACTGGTTCAATCCAAAGACGCCCGGATGTTTTCAGAGTACACCCG 660
DB	601 cctcagcta tgggactggttcaatccaaagaaacccgaggtgttttgacagtgaccctg 660
QY	661 TGGAAAGCCCGATTTGTGGGAGGACATTATCACACAGCTCTGCTGGAAAGTACTATC 720
DB	661 tggaaaggccgattgtgtgggaagcaacttatgacacagctctgctggaaagtactac 720
QY	721 GCCACACAGAAACTCAGTGTGGGCTGACAGTGTTCCTGTGGGAAAGTACATTTGAGCAT 780
DB	721 gccacacagaaactcactgtggggctgacagtgcttctgttgggaaagtacatttgagcat 780
QY	781 TACTTTAGAACACTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGTCATA 840
DB	781 tactttgaagacctctctggagctctgacatgtactctcatggttggccatctgggtcata 840
QY	841 TTTTACGCTCATGATAGACGACACTCCCGGATGCCTGCTGACACTGAACCCCTCTACAT 900
DB	841 ttttagctcatgatagacgacacctcccgatgctgtcgtgcacctgaaacctctacat 900
QY	901 TCCTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGTGGCAGGATATCAGCATGATCGCG 960
DB	901 tccttcaagctcttctgagatcaggtctgagaagaggttggcaggatatacagcatgatcgc 960
QY	961 ATGAAGACCAATTGGGAGACATCTCGGCCACATTCAGCACGAGTTCGACTTCCTCTTC 1020
DB	961 atgaagaccaattggggagacacctctggccacacatccagcaaggtcagacttctcttc 1020
QY	1021 TGATGACGCTGGATCAAGTCTTTCAAGACAATTCGGGGTGGAAACTCTGGGCCAGCTG 1080
DB	1021 tgcattgagctggatcaagtcttccaagacaactctggggtggaaactctgggccaactg 1080
QY	1081 GTAGCAGCTCCAGGCTCGTGGTACAAGGCCAGTCCCGAGAGTTCACCTATGAGAG 1140
DB	1081 gtgacagctccaggcctggtgtacaaggccagtcctccgagaagttcacctatgagagg 1140
QY	1141 CGGGAACCTGTGGCCCGGTACATTCCATTCGAGAGGGGATTTTACTACCAACGCGGCC 1200
DB	1141 cgggaactgtcggccgcgtacattccattcggagaggggagatttttactacacagcgcc 1200
QY	1201 ATTTTGGAGAACGCTTACTACATTTCTCAACTCTACAGGAGTGTCTTAAAGGGGATC 1260
DB	1201 attttggaggaaacgcttactcacattctcaacctcacccagggagtgctttaaggggatc 1260
QY	1261 CTCAGGACAAGAAACATGACATAGAACCCAGTGGCATGATGAGACCCACCTCAACAAA 1320
DB	1261 ctccaggacaagaaaatgacatagaagcccatgtggcatgtgagagccaccctcaacaaa 1320
QY	1321 TACTTCTTTTCAACAAACCCACTAAAATCCTATCTCCAGAGTATTTGCTGGGACTATPAC 1380
DB	1321 tacttctctttcaacaaacccactaaaaatcctatctccagagatttgcgtgggactatcag 1380
QY	1381 ATAGGCTCGCTTTCAGATATTAAGTGTCAAGTACGTTGGCAGACAAAGAGTATAAT 1440
DB	1381 ataggctcgcttccagatattaaaagtgtcaaggtagcttggcagacaaaagagataaat 1440
QY	1441 TTGTTTAGAATAATATCTCTGACTTCAAAATTTGATGAAACTTCACACTATTTCTAACCA 1500
DB	1441 ttggttagaataatattctctgacttcaaaatgtgatgaaacttcacactattttcttaacca 1500

Db 1441 ttggttagaataatgtctgacttcaaatgtgatgaaacttgacactatttctaacca 1500

RESULT 4  
AAD26484  
ID AAD26484 standard; DNA; 1500 BP.  
XX  
AC AAD26484;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Mouse alpha (1, 3) galactosyltransferase DNA.  
XX  
KW Mouse; xenotransplantation; Gal-alpha (1,3)Gal; GAL determinant; surgery;  
KW alpha(1,3) galactosyltransferase; alpha (1, 3) GT; drug screening;  
KW gene therapy; ds.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 277..1461  
FT /tag= a  
FT /product= "Mouse alpha (1, 3) GT protein"

WO200188096-A2.  
XX  
PD 22-NOV-2001.  
XX  
PF 14-MAY-2001; 2001WO-US15765.  
XX  
PR 15-MAY-2000; 2000US-204148P.  
PR 13-JUN-2000; 2000US-0593316.  
XX  
PA (GERO-) GERON CORP.  
XX  
PI Denning C, Clark J;  
XX WPI; 2002-089848/12.  
DR P-PSDB; AAE16321.  
XX  
PT New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal  
PT determinants, for xenotransplantation, and in the treatment of the  
PT human body by surgery or therapy -  
XX  
PS Claim 16; Page 67-68; 86pp; English.  
XX  
CC The patent discloses immunologically compatible animal tissue, suitable  
CC for xenotransplantation into human patients. The invention also relates  
CC to an ovine tissue devoid of antibody-detectable Gal-alpha (1,3)Gal (GAL)  
CC determinants which are made by alpha(1,3) galactosyltransferase (GT). The  
CC ovine tissue is useful for treatment of human body by surgery or therapy  
CC and in xenotransplantation, by transplanting the ovine tissue into a  
CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.  
CC Polynucleotide constructs of the invention are useful for inactivating  
CC an alpha(1,3) GT gene in an ovine cell. Alpha (1,3) GT sequences are useful  
CC for drug screening and for the production of GAL containing synthetic  
CC oligosaccharides. Sequences of the invention are also useful in gene  
CC therapy. The present sequence is a DNA encoding mouse alpha (1,3) GT  
XX protein.  
XX  
SQ Sequence 1500 BP; 406 A; 349 C; 374 G; 371 T; 0 other;

Query Match 100.0%; Score 1500; DB 24; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CCTTCCCTTGTAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCCGCGT 60  
Db 1 ccttccctttagactcttcttggaaatgagaagtaccgattctgctgaagacctcgctg 60  
Qy 61 CTGAGGCTCTGGAGTTGGAAACCTGTACCTTCTCTCTCTGCTGAGCCCTGCTTCTT 120  
Db ccttccctttagactcttcttggaaatgagaagtaccgattctgctgaagacctcgctg 120

Db 61 ctgaggtctgagggttgaaacctgtaccttcccttctctctgctgagccctgcctctt 120  
Qy 121 AGGCAGCCAGAGCTCGACAGAACTCGGTTTGGCTTTCGCTTTCGCTTTCGAGGCAACACAG 180  
Db aggcagggccagagcttcgacagaaactcggctgcttctgctcttctgaggggaaacacag 180  
Qy 181 CTGACGATGAGGCTGACTTTTGAACTCAAGAGATCTGCTTACCCCCAGTCTCTCTGGAATTA 240  
Db ctgacgatgaggctgacttgaactcaagagactgcttaccaccagctcctctggaattaa 240  
Qy 241 AGGCCTGTACTACATTTGCCTGACCTAAGATTTTTCATGATCATTGCTTCAAGATCTC 300  
Db aggcctgtactacatttgcctggaccataagatttcatgatcactatgcttcaagactc 300  
Qy 301 CATGTCAACAAGATCTCCATGTCAAGTCCAAGTCAAGAACTGCTTCCATCCTCAAGA 360  
Db catgtcaacaagatctccatgtcaagatccaaagtcagaacaagctctccatcctcaaga 360  
Qy 361 TCTGGATCACAGAGAAATAATGAATGTCAAGGAAAGAACTAATCTCTGTGTGCTGTGATT 420  
Db tctggatcacagagagaaaaataatgaatgtcaagggaaaaagtaactcctgttgaagtgt 420  
Qy 421 GTCTCAACCGTGTGTTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG 480  
Db gtctcaaccgt 480  
Qy 481 AACAGATGGCAGAGGACTGCTGTTTCCCAAGCTGGTTTAAATAATGGAGCCACAGTTAT 540  
Db aacagatggcagagagactgt 540  
Qy 541 CAAGAAGACAACTAGAGGAGGAGAGAAAGGGTAGAAATCGAGATCCATTTGAAGAG 600  
Db caagaagacaactag 600  
Qy 601 CCTCAGCTATGGGACTGCTTCAATCCAAAGACCGCCCGGATGTTTTCAGACGTGACCCG 660  
Db cctcagctatgggactgt 660  
Qy 661 TGAAGGCGCGGATTTGTGGGAGGACACTTATGACACAGCTCTGCTGGAAAAGTACTATC 720  
Db tgaaggcgccgattgt 720  
Qy 721 GCCACAGAAACTCCTACTGTGGGGCTGACAGTGTGCTGTGGGAAAGTACATTTGAGCAT 780  
Db gccacagaaaactcactgt 780  
Qy 781 TACTTACAGACATTTCTGGAGTCTGTGACATGTACTTTCATGTTGCCCATCGGCTCATA 840  
Db tacttgaagacttcttggagtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840  
Qy 841 TTTTACGTCTATGACAGACACTCTCCGGATGCTTCTGTCGACCTTGAACCTCTTACAT 900  
Db ttttaagctatgataagacacactcctccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 900  
Qy 901 TCTTTACAACTCTTTGAGATCAGGTCTGAGAAGAGGTGGCAGGATATCAGCATGATCGGC 960  
Db tcttcaagctctttagatcaggctcagagagagagagagagagagagagagagagagagag 960  
Qy 961 ATGAAGACCATTCGGGAGGACATCTCCGCCACATCCAGCAGCGAGGTGCTTCTCTTTC 1020  
Db atgaagaccatttgggagagacactccttggccacatccagcacaggttcgacttctcttct 1020  
Qy 1021 TGCATGACCTGGATCAAGTCTTTTCAAGACAACTTCCGGGTGGAAGTCTTGGGCCAGCTG 1080  
Db tgcattgagctggatcaagcttcttcaagacaaacttcggggttggaacctctgggcccagctg 1080  
Qy 1081 GTAGCAGCTCCAGGCTTGGTGTACAGGCGAGTCCCGAGAGTTTTCACCTATGAGAGG 1140  
Db gtacacagctccaggtcgt 1140  
Qy 1141 CGGGAACCTGCGGCCGCTACATTCATTCCGAGAGGAGGGGATTTTACTACCAGCGCGCC 1200  
Db cgggaactgtcggccgctgacatttccaattcggagagggggttttttactaccacgcggcc 1200

Qy	1201	ATTTTGGAGGAACGCTACTCACATTTCTCAACCTCACCAGGAGTGCTTTTAAGGGGATC	1260
Db	1201	attttggagggaacgcctactcacattctcaacctccaccaggagtgcttttaaggggatc	1260
Qy	1261	CTCAGGACAAGAAACATGACATAGAACCCAGTCGGCATGATGAGAGCCACCTCAACAAA	1320
Db	1261	ctcaggaacaagaaacatgacatagaagcccgatggcatgatgagccaccctcaacaaa	1320
Qy	1321	TACTTCTCTTTTCAACAAACCCACTAAAATCTTATCTCAGAGTATTGCTGGGACTATCAG	1380
Db	1321	tacttctcttttcaacaacccactaaaaatcctatctccagagtatgtctgggactatcag	1380
Qy	1381	ATAGGCTGCTTTCAGATATTAAAGTGTCAAGGTAGCTTCGCAGACAAAGAGTATAAT	1440
Db	1381	ataggctgccttcagatatataaagtgctaaaggtagctctggcagacaaagagcaaat	1440
Qy	1441	TTGGTTAGAAATAATGCTGTGACTTCAAAATGTGTATGGAAACTTGCACACTATTTCCTAACCA	1500
Db	1441	tggtttagaaaataatgtctgcacttcaaatgtgtagtggaaacttgacactatttcttaacca	1500

CC consisting essentially of transgenic cells engineered, for use in  
CC transplantation. The methods of the invention facilitate  
CC xenotransplantation between species, particularly between species  
CC exhibiting differential expression of the gal-alpha-gal epitope. The  
CC invention also describes primers (AAS05334-AAS05385) used to isolate the  
CC alpha-1,3 GT sequences.  
XX  
XX Sequence 3537 BP; 978 A; 777 C; 843 G; 939 T; 0 other;

QY 864 CTCCTCCGATGCTCTGCTGCACCTGAACCTCTTACATCTCTTACAGTCTTTGAGATCAG 923  
 DB 1020 CTCCTCGATGCTCTGCTGCACCTGAACCTCTTACATCTCTTACAGTCTTTGAGATCAG 1079  
 QY 924 GTCTGAGAAAGAGTGGCAGGATATCAGCATGATGCCCATGAAGACCATTTGGGAGGACAT 983  
 DB 1080 GTCGAGAGAGGTGGCAGGATATCAGCATGATGCCCATGAAGACCATTTGGGAGGACAT 1139  
 QY 984 CTGGGCCACATCCACGACGAGTGCACCTTCTCTCTGATGACCTGCATCAAGCTCTT 1043  
 DB 1140 CTTGGCCACATCCACGACGAGTGCACCTTCTCTCTGATGACCTGCATCAAGCTCTT 1199  
 QY 1044 TCAGAACACTTTCGGGGTGGAACTCTGGGCCAGCTGGTAGCACACCTCCAGGCCCTGGTG 1103  
 DB 1200 TCAAGAACACTTTCGGGGTGGAACTCTGGGCCAGCTGGTAGCACACCTCCAGGCCCTGGTG 1259  
 QY 1104 GTACAGGCCAGTCCCGACGAGTTCACCTATGAGAGCGGGAACTCTCGGCCCGTACAT 1163  
 DB 1260 GTACAGGCCAGTCCCGACGAGTTCACCTATGAGAGCGGGAACTCTCGGCCCGTACAT 1319  
 QY 1164 TCATTCGGAGAGGGGATTTTACCTACCCAGCGGCTATTTTGGAGAACGCTACTCA 1223  
 DB 1320 TCCATTCGGAGAGGGGATTTTACCTACCCAGCGGCTATTTTGGAGAACGCTACTCA 1379  
 QY 1224 CATCTCAACCTCACAGGAGTGTCTTTAAGGGATCTCTCCAGGACAGAAACATGACAT 1283  
 DB 1380 CATCTCAACCTCACAGGAGTGTCTTTAAGGGATCTCTCCAGGACAGAAACATGACAT 1439  
 QY 1284 AGAAGCCAGTGCATGATGAGAGCCACCTCAACAATACTTCTTTCACAAACCCAC 1343  
 DB 1440 AGAAGCCAGTGCATGATGAGAGCCACCTCAACAATACTTCTTTCACAAACCCAC 1499  
 QY 1344 TAAATCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCTCTCAGATATTA 1403  
 DB 1500 TAAATCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCTCTCAGATATTA 1559  
 QY 1404 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATATTTGGTTAGAAATAATCTCTGACT 1463  
 DB 1560 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATATTTGGTTAGAAATAATCTCTGACT 1619  
 QY 1464 TCAAAATGTGATGGAACATGTGACACTATTCT 1495  
 DB 1620 TCAAAATGTGATGGAACATGTGACACTATTCT 1651  
 RESULT 6  
 AAT04523  
 ID AAT04523 standard; cDNA; 1116 BP.  
 AC AAT04523;  
 AT AAT04523;  
 DT 01-MAY-1996 (first entry)  
 XX Mouse alpha-1,3-galactosyltransferase coding region.  
 DE Mouse alpha-1,3-galactosyltransferase coding region.  
 XX Mouse; alpha-1,3-galactosyltransferase; Immune response; glycoprotein;  
 KW alpha-galactosyl epitope; cell membrane; virus; phagocytosis; tumour;  
 KW antigen processing; leukemia; lymphoma; myeloma; melanoma; carcinoma;  
 KW sarcoma; vaccine; opsonisation; glycoprotein; antibody; anti-Gal; ss.  
 XX Mus musculus.  
 OS Mus musculus.  
 XX WO9524924-A1.  
 PN 21-SEP-1995.  
 PD 13-MAR-1995; 95WO-US03156.  
 XX 15-MAR-1994; 94US-0213200.  
 PR (UYHA-) UNIV HAHNEMANN & MEDICAL COLLEGE PENNSYL.  
 PA (UYHA-) UNIV HAHNEMANN & MEDICAL COLLEGE PENNSYL.  
 XX

PI Gallili U, Repik PM;  
 XX WPI; 1995-336816/43.  
 DR Association of an alpha-galactosyl epitope with a tumour or viral  
 XX antigen - is administered to anti-Gal synthesising animals to induce  
 PT an immune response  
 PT Disclosure; Fig 9; 85pp; English.  
 PS The nucleotide sequence of the mouse alpha-1,3-galactosyltransferase  
 XX coding sequence. The enzyme can be used in methods of enhancing an  
 CC immune response by associating the alpha-galactosyl epitope with a cell  
 CC membrane or viral glycoprotein. The alpha-galactosyl epitope enhances  
 CC phagocytosis and subsequent processing of the antigen. The method is  
 CC useful in the treatment of tumours e.g. leukemia, lymphoma, myeloma,  
 CC melanoma, carcinoma and sarcoma, or for the generation of viral vaccines  
 CC by opsonising a viral glycoprotein. The alpha-galactosyl epitope  
 CC enhances recognition of the antigen in an animal that synthesises the  
 CC naturally occurring antibody - anti-Gal.  
 CC See AAT04522 for the marmoset alpha-1,3-galactosyltransferase sequence.  
 XX Sequence 1116 BP; 309 A; 250 C; 293 G; 264 T; 0 other;  
 SQ  
 Query Match 68.7%; Score 1030.8; DB 16; Length 1116;  
 Best Local Similarity 96.6%; Pred. No. 3e-310;  
 Matches 1078; Conservative 0; Mismatches 2; Indels 36; Gaps 1;  
 QY 382 ATGAATGTCAAGGAAAGTAATCTCTGTGATCTCTCAACCGTGTGCTGCTG 441  
 DB 1 atgaatgtcaaggaaagtatctgtgtgatctgtatgtctcaacccgtgtgtctgtg 60  
 QY 442 TTTTGGGATATCTCAACAG-----AATT 465  
 DB 61 ttttgggaatatctcaacagccagagcgtcttcttgggtatatacaacaaatt 120  
 QY 466 CCAGAGTGTGTGAGAACAGATGGCAGAGGACTGTGTGTTCCCAAGCTGTTTAAAT 525  
 DB 121 ccagaggttggtgagacagatggcagagagactggtgttcccaagctggtttaaatt 180  
 QY 526 GGGACCCACAGTTATCAAGAGACACGTTAGAGGACGGAGAGAAAGGTCAGATGGA 585  
 DB 181 gggacccacagtatcaagaagacaacgtagaaggcggagagaagaggtagaattgga 240  
 QY 586 GATCGCATTTCAAGAGCTCTAGCTATGGGACTGTCTCAATCCAAAGACCCCGGATGTT 645  
 DB 241 gatcgcattgaagagcctcagctatgtggactggttcaatccaaagaaccccgagatgtt 300  
 QY 646 TTGACAGTACCCCGTGGAGGCCGATTTGTGTGGGAGGCACATTATGACAGCTCTG 705  
 DB 301 ttgacagtgaccccggtggaagcgccgattgtgtggaagggcacttatgacacagctctg 360  
 QY 706 CTGGAAGTACTACGCCACACAGAACTCTGTGGGCTGACAGTGTGCTGCTGCGGA 765  
 DB 361 ctggaagtgactacgcccacacagaaactcactgtgggctgacagtgctgtgtgtgga 420  
 QY 766 AACTACATTTGAGCATTTAGAAAGCTTTCTCGAGTCTGCTGACATGTACTTTCATGTT 825  
 DB 421 aactacatgtgagcatctacttagaagacttctctgtgagcctgtgctgacatgtacttca 480  
 QY 826 GGCCATCGGTCATATTTTACGTATGATAGACACACCTCCCGGATGCTGCTGTCAC 885  
 DB 481 ggccatcggttcataatttttaactcagtcagtagacacacctcccgatgctgtcgtgac 540  
 QY 886 CTGAACCCCTCTACATCTTACAGCTTTTGGAGATCAGGTCCTGAGAGAGGTCGCCAGAT 945  
 DB 541 ctgaacctctacatctccttcaagctctttgagatcaggtctgagaagaggtgagcagat 600  
 QY 946 ATCAGCATGTGCGCATGAAGACCATTTGGGAGCAGCATCTCTGCCCATCTCCAGCAG 1005  
 DB 601 atcagcatgtgctgcatgaagaccattggggagcacatctctggtcccaacatccacgacgag 660

QY 1006 GTCGACTCTCTCTTCGATGGACGCTGGATCAAGCTCTTCAAGACAACTTCGGGGTGGAA 1065  
Db 661 gtgactctctctctgcatgagctggatcaagctcttcaagacaactctcgsggtggaa 720  
QY 1066 ACTCTGGCCAGCTGGTAGCACAGCTCCAGGCTGGTGTACAAAGCCAGCTCCCGAGAAG 1125  
Db 721 actctggccagctcggtagcacagctccaggctcggtagcaaggccagctcccgagaac 780  
QY 1126 TTCACCTATGAGAGCGGGGAACCTCGCGCGGTACATTCATTCGGAGAGGGGATTTT 1185  
Db 781 ttcacatagagagcggaactgtcgccgctacatctccattccgagaggggatttt 840  
QY 1186 TACTACCACGGCCATTTTGGAGAACGCTACTCACATTTCAACCTCACAGGAG 1245  
Db 841 tactaccacgcgccattttggaggagcctactcacattctcaacctcaccaggag 900  
QY 1246 TGCTTTAAGGGATCTCCAGGACAACAAATGACATAGAACCCAGTGCATGATGAG 1305  
Db 901 tgccttaagg99ga tccctccg99gaagaacatgacatagaagcccg9tgcata 960  
QY 1306 AGCCACCTCAACAAATCTCTCTTTTCAACAAACCCACTAAAATCCTATCTCCAGAGTAT 1365  
Db 961 agccacctcaacaatactctcttttcaacaacccactaaaatcctatctccagagtat 1020  
QY 1366 TGCTGGGACTATCAGATAGGCTGCCCTTCAGATATTTAAAAGTGTCAAGGTAGCTTGGCAG 1425  
Db 1021 tgcctggactatcagatagcctgccttcagatatataaagtgtcaaggtagcttggcag 1080  
QY 1426 ACAAAAGATATATTTGGTTAGAAATAATGCTCGA 1461  
Db 1081 acaaaagatataatttggttagaaaataatgctcga 1116

## RESULT 7

AAQ74712  
ID AAQ74712 standard; cDNA to mRNA; 1423 BP.  
XX  
AC AAQ74712:  
XX  
XX  
DT 26-JUN-1995 (first entry)  
XX  
DE Galactosyl transferase clone.  
XX  
KW Gal-alpha (1,3) galactosyl transferase; xenograft; transplant;  
KW rejection; ss.  
XX  
XX Sus scrofa domestica.  
XX  
XX Key Location/Qualifiers  
FH CDS 50..1129  
FT /\*tag= a  
XX  
XX WO9421799-A.  
XX  
XX 29-SEP-1994.  
XX  
XX 15-MAR-1994; 94WO-A000126.  
XX  
XX 16-MAR-1993; 93AU-0007854.  
XX  
XX (AUST-) AUSTIN RES INST.  
XX  
XX Mckenzie IFC, Sandrin MS;  
XX  
XX WPI; 1994-317019/39.  
XX  
XX P-PSDB; AAR62508.  
XX  
XX DNA sequences encoding Gal-alpha (1,3)galactosyl transferase -  
XX and clones contg. such sequences are used in xenograft therapies  
XX  
XX Claim 3; Page 35; 50pp; English.  
XX  
XX The sequence is that of the porcine Gal-alpha (1,3) galactosyl

CC transferase gene which produces a Gal epitope on the surface of  
CC porcine cells. This epitope is recognised by antibodies which are  
CC responsible for hyperacute rejection of xenotransplanted pig cells,  
CC tissues and organs  
CC See also AAQ74711-4.  
XX  
SQ Sequence 1423 BP; 413 A; 305 C; 348 G; 357 T; 0 other;

Query Match 48.4%; Score 726.6; DB 15; Length 1423;  
Best Local Similarity 78.8%; Pred. No. 2.1e-215;  
Matches 893; Conservative 0; Mismatches 234; Indels 6; Gaps 2;

QY 354 CTCGAAGTCTGGATCACAGGAGAAAATAATGAATGTCAAGGGAAGTAATCTCTGTGAT 413  
Db 22 cccagctctcgccatcaggagaaaaaataatgaatgtcaagggaagagtggtctcgtcaat 81  
QY 414 GCTGATTTGCTCAACCGTGGTTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGT 473  
Db 82 gctgcttgctcaactgtaatggttggtttgggaatacatcaacagaaacccagagagt 141  
QY 474 TGGTGAAACAGATGGCAGAGGACTGGTGTCCCAAGCTGGTTTAAAATGGAGCCCA 533  
Db 142 tgg ---cagcagtgctcagaggggctggtggtttcccgagctggtttcaaatgggactca 198  
QY 534 CAGTTATCAAGAAGACAACCTAGAAAGCAGGAGAGAAAAGG ---GTAGAAATGGAGATCG 590  
Db 199 cagttaccagagaagaagacgctataggcaacgaaaaggaaacaaagaaaagaagacaa 258  
QY 591 CATTGAAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCGCGGATGTTTGAC 650  
Db 259 cagaggagagcttcgcgtagtggactggtttaactcctgagaacgcccagaggtcgtgac 318  
QY 651 AGTGACCCCGTGGAAAGCGCGGATTTGTGGGAAGGCACCTTATGACACAGCTCTGCTGA 710  
Db 319 cataaccagatgggaaggtccagtggtatggaaggacattcaaacagagcgctctaga 378  
QY 711 AAAGTACTACGCCACACAGAAACTCACTGTGGGCTGCACAGTGTGTTGCTGTGGAAAAGTA 770  
Db 379 taattattatgccaaacagaaaaattaccgtgggcttgagcgttttctgctgcggaagata 438  
QY 771 CATTGACCATTTACTTAGAAGACTTTCTGGAGTCTGCTGACATGTACTTCATGTTGGCCA 830  
Db 439 cattgagcatctactggagaggtcttcaatctgcaaatatactactcatggttggcca 498  
QY 831 TCGGTCATATTTTACGTCATGATGACAGCACACCTCCGGATGCTGCTGTCACCTGAA 890  
Db 499 caaagtcatctttacatcatcgtggatgatactccaggatgctcttgatagagctggg 558  
QY 891 CCCTCTACATTCCTTACAAAGTCTTTTGAGATCAGGTCTGAGAAGAGGTGGCAGGATATCAG 950  
Db 559 tcctctgcttctcttaagtggttgagatacaagtcgagagaagaggtggcaagacatcag 618  
QY 951 CATGATGGCATGAAGACCATTTGGGGAGCACATCTCGCCACATCCAGCAGAGGTGCA 1010  
Db 619 catgatcgcatgaagaccatcgggagacatctctggccacatccagcagcaggtggga 678  
QY 1011 CTTCTCTTCTGCAATGGACGTGAAGTCAAGTCTTTCAAGACAACCTTCGGGGTGGAACTCT 1070  
Db 679 ctctctcttgcattgacgtggatcaggtcttccaaaacaactttggggtggagaccct 738  
QY 1071 GGGCCAGCTGGTAGCAGCTCCAGGCTGGTGTGTACAAAGCCAGTCCCGAGAAGTTTAC 1130  
Db 739 gggccagtcggtggctcagctacagggcgtggtggtatacaggcacatcctcagcaggttcac 798  
QY 1131 CTATGAGAGCGGGAACCTGTCCGCCGCTACATTCATTCGAGAGGGGATTTTACTA 1190  
Db 799 ctacgagagcggaagagtcgcgagcctacattccgtttggccagggggtattttatta 858  
QY 1191 CCAGCGCGCCATTTTGGAGGAAGCCCTACTTCATATCTCAACCTCACCAGGGAGTGT 1250  
Db 859 ccagcagccatttttgggggaacacccactcaggttcttaaacatcactcagaggtgctt 918



QY 1251 TAAGGGATCTCCAGGACAAGAACATGACATAGAGCCCGAGTGGCATGATCAGAGCCA 1310  
DB 919 caagggaatcctccaggacaaggaataatgacatagaaagccgagtggaatgaaagcca 978  
QY 1311 CCTCAACAATACTTCCCTTTTCAACAAACCCACTAAATCCTATCTCCAGAGTATTGCTG 1370  
DB 979 tctaacaagatattcctctcaacaacccactaaaaatctatacccagaatactgctg 1038  
QY 1371 GGACTATCAGATAGGCGCTGCCCTTCAGATATTAAAGTGTCAAAGGTAGCTTTGGCAGACAAA 1430  
DB 1039 ggaattatcatagcatgctctgtggatattaggattgccaagatagcttggcagaataa 1098  
QY 1431 AGAGTAAATTTGGTTAGAAATAATGCTCTGACTTCAAAATTGATGGAAACATT 1483  
DB 1099 agagtataattgggttagaataaacatctgacttttaaatgtgcccagcagttt 1151

RESULT 8

AAT12242  
ID AAT12242 standard; cDNA to mRNA; 1423 BP.

AC AAT12242;

DT 08-APR-1996 (first entry)

DE Pig alpha(1,3)-galactosyltransferase, full coding sequence.

XX Alpha(1,3)galactosyltransferase; xenograft hyperacute rejection;  
KW transplantation; galactose alpha(1,3) galactose; ds.

XX Sus scrofa.

XX Key Location/Qualifiers  
FT CDS 50..1129  
FT /\*tag= a

XX W09534202-AL.

XX 21-DEC-1995.

XX 14-JUN-1995; 95WO-US07554.

XX 21-JUL-1994; 94US-0278282.

PR 15-JUN-1994; 94US-0260201.

XX (ALEX-) ALEXION PHARM INC.

PA (AUST-) AUSTIN RES INST.

PI Fodor WL, McKenzie IFC, Rother RP, Sandrin MS, Suinto SP;

XX WPI; 1996-049326/05.

DR P-PSDB; AAR90573.

XX Redn. of rejection of xenogenetic cells following transplantation  
PT by introducing a vector expressing fucosyl:transferase into the  
PT cells

XX Example 2; Page 52-54; 69pp; English.

XX The cDNA (AAT12242) encoding pig alpha(1,3)-galactosyltransferase  
CC (AAR90573) was cloned using cross species hybridization. Vector  
CC pGT contains an insert comprising this cDNA sequence under the  
CC control of the cytomegalovirus promoter. The vector was  
CC co-transfected with vector pHT encoding human H-transferase  
CC (AAR90572) into COS cells. This resulted in a reduction in the  
CC levels of galactose alpha(1,3) galactose epitopes expressed by  
CC the cells.

XX Sequence 1423 BP; 413 A; 305 C; 348 G; 357 T; 0 other;

Query Match

Best Local Similarity 48.4%; Score 726.6; DB 17; Length 1423;

78.8%; Pred. No. 2.1e-215;

	Matches	893;	Conservative	0;	Mismatches	234;	Indels	6;	Gaps	2;
QY	354	CTCAAGATCTGGATCACAGGAGAAATAATGAATGTCAAGGGAAAGTAATCTCTGTGAT	413							
DB	22	cccagctctgcgcgaccaggagaaaaataatgaatgccaaggaagagtggtctctgccaat	81							
QY	414	GCTGATGTCTCAACCGTGTCTGCTGTTTGGGAATATGTCAACAGAAATTTCCAGAGGT	473							
DB	82	gctgctctcacaactgtaatggttgctgttttggaatacatcaacacagaccagagtg	141							
QY	474	TGCTGAGAACACATGGCAGAGGACTGGTGTCTCCCAAGCTGTTTAAAAATGGGACCCA	533							
DB	142	ttg --- cagcagctgctcagagggctgggttggtttccgagctggtttacaatggggacta	198							
QY	534	CAGTTATCAAGAAGCAACGTTAGAAGCGGAGAGAAAAAGG---GTAGAAATGGAGATCG	590							
DB	199	cagttaccacgaagaagacgcctataggcaagaaaaaggaacaaagaaagaaagacaa	258							
QY	591	CATTGAAGAGCCTCAGCTATGGGACTGTGTTCAATCCAAAGAACCCGCCGATGTTTGTAC	650							
DB	259	cagaggagagctccgctagtggaactggtttaactctgagaaacgccagagtcgtgac	318							
QY	651	AGTGACCCCGTGGAAAGCGCCGATGCTGTGGGAAGCACCTTATGACACAGCTCTGCTGGA	710							
DB	319	cataaccagatggaaggtccagtggtatcggaaggaacttaacacagagccgtctaga	378							
QY	711	AAAGTACTAGCCACACAGAAACTCACTGTGGGCTGACAGTGTCTTGTGGGAAGTA	770							
DB	379	taattattatgcaaacagaaaaattaccgtgggcttgacgggttttctgctgcggagaata	438							
QY	771	CATTGAGCATTTACTTAGAAGACTTTCTTGGAGTGTGCTGACATGTACTTTCATGGTGGCCA	830							
DB	439	cattgagcattacttgaggaggtcttataatactgcaaatatacttcatgttggtgccaa	498							
QY	831	TGCGGTTCATATTTTACGTCATGATAGACGACACCTCCCGGATGCCCTGTGCTGCACCTGAA	890							
DB	499	caaaagtcattcttatcatcatggtggtatgatatctccaggtgcgtttgatagagctggg	558							
QY	891	CCCTCTACATTTCTTACAAGTCTTTTTCAGATCAGGTCTGAGAAGAGTGGCAGGATATCAG	950							
DB	559	tectctgcttctcttaaaagtgtttgagatcaagtcaggagaaagagtggaagacatcag	618							
QY	951	CATGATGGCATGAAGACCATTTGGGAGCACATCTCTGGCCACATCCAGCACAGAGTCCA	1010							
DB	619	catgatgcgatgaagaccatcggggagcacatccttgcccacatccagcacgagtgga	678							
QY	1011	CTTCTCTTCTGCATGGAGCTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAAACTCT	1070							
DB	679	cttctctctctgcatgtgacgtggtatcagctcttccaaaaaacactttgggggtggagaccct	738							
QY	1071	GGGCCAGCTGGTAGCACAGCTCCAGGCCCTGGGTGTACAAGGCCAGTCCCGAGAGTTTCAC	1130							
DB	739	gggcccagtcgggtgggtccagctacagggccttgggtggaaggggacacacctgacagtgccac	798							
QY	1131	CTATGAGAGCGGGAACTGTCTGGGCCCGCTACATTTCCATTTCCGAGAGGGGATTTTACTTA	1190							
DB	799	ctacgagagcggaagagagtcgcgcagcctacatcttcgtttggccagggggttttatca	858							
QY	1191	CCACGGCGGCATTTTGGAGGAACGGCTACTCATATTTCTCAACCTCACCAGGAGTGCCTT	1250							
DB	859	ccacgcagcctttttgggggaacacccactcaggttcttaaacatcactcagagtgctt	918							
QY	1251	TAAAGGGATCTCCAGGACAAAGAACATGACATAGAGCCCGACGTGCGATCATGACAGACCA	1310							
DB	919	caagggaalccctccaggacaaggaataatgacaagaagcgagtggaatgaaagacca	978							
QY	1311	CCTCAACAAATACTTCTCTTTTCAACAAACCCACTAAAAATCCTATCTCCAGAGTATTGCTG	1370							
DB	979	tctaacaagatattctcttctcaacaacccactaaaaatcttatccccagaatactgctg	1038							
QY	1371	GGACTATCAGATAGGCGCTTCGCTTTCAGATATTAAAGTGTCAAGGTAGCTTGGCAGACAAA	1430							
DB	1039	ggattatcataaggcatgtctctggatattaggttgtaagattgtcaagatagcttggcagaaaa	1098							



DE Porcine alpha-1,3-galactosyl transferase isoform 2 cDNA.  
XX  
KW Isoform: porcine; enzyme: alpha-1,3-galactosyl transferase; galactose;  
KW sugar; N-acetyllactosamine; glycoprotein; glycolipid; antibody; pig;  
KW graft tissue rejection; organ transplantation; xenotransplant; ss.  
XX  
QS Sus scrofa.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1090  
FT /tag= a  
FT /product= "alpha-1,3-galactosyl transferase isoform 2"  
FT /note= "contains no start or stop codon at the 5' or 3'  
FT ends of the sequence"  
XX  
PN FR2751346-A1.  
XX  
PD 23-JAN-1998.  
XX  
PF 19-JUL-1996; 96FR-0009077.  
XX  
PR 19-JUL-1996; 96FR-0009077.  
XX  
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
XX  
PI Pourcel C, Soullillon JP, Vanhove B;  
XX  
DR WPI; 1998-112876/11.  
XX  
DR P-PSDB; AAW49687.  
XX  
PT Transgenic non-human donors of organs for human recipients -  
PT containing DNA encoding antibodies that inhibit graft rejection  
XX  
PS Claim 4; Page 36-37; 71pp; French.  
XX  
CC This sequence represents the cDNA encoding isoform 2 of the porcine  
CC enzyme alpha-1,3-galactosyl transferase (alpha-1,3-GT). The enzyme  
CC catalyses the attachment of a galactose sugar molecule on the  
CC N-acetyllactosamine moiety found on surface glycoproteins and  
CC glycolipids. These sugar molecules are partly responsible for raising  
CC anti-graft antibodies, which lead to graft tissue rejection. The  
CC invention relates to a method of inhibiting the graft rejection mechanism  
CC by introducing the sequence encoding an antibody targeted to alpha-1,3-GT  
CC into the cells of animal, especially a pig, from whom organs may be used  
CC for xenotransplants. Neutralisation of the alpha-1,3-GT leads to tissues  
CC or organs lacking the galactose on the glycoproteins and glycolipids,  
CC thus preventing induction of the rejection response.  
SQ Sequence 1092 BP; 325 A; 228 C; 279 G; 260 T; 0 other;

Query Match 47.1%; Score 706; DB 19; Length 1092;  
Best Local Similarity 79.3%; Pred. No. 4.6e-209;  
Matches 864; Conservative 0; Mismatches 220; Indels 6; Gaps 2;

QY 375 GAAATAATGAATGTCAAGGAAAGTAATCTGTGTGATGTCTCAACCGTGTG 434  
Db 6 gaaaaLaatgaatgtcaaggaagagtggtctctcaatgctgctgtctcaactgaat 65  
QY 435 TGTCTGTGTTTCGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAGAACAGATGGCGAA 494  
Db 66 ggtctgtgttttgggaatacatcaacagaaaccagaaagtgtg---cagcagtgctcagag 122  
QY 495 GGACTGCTGTTCCTCCCAAGCTGGTTTAAAAATGGAGCCACACAGTTATCAAGAGACAAACGT 554  
Db 123 gggctgtgtgtttccagagctgtttaaataatgggactcaacagttaccacagaagaaga 182  
QY 555 AGAAGACGCGAGAGAAAAG---GTAGAAATGGAGATCCGATTTGAAGACCCCTCAGCTATG 611  
Db 183 cgtataggaacagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 242  
QY 612 GGACTGTTTCATCCAAAACCGCCCGGATGTTTTCAGACATGACCCCGTGGAGCGGCC 671  
XX

Db 243 ggaactggtttaatactctgagaaagccagaggtcgtgaccaaaacagatggaaggtccc 302  
QY 672 GATTGTGTGGGAGGACACTTATGACACACAGCTCTCTGGGAAAAGTACTACGCCACACAGAA 731  
Db 303 agtgataggaaggaagcacttacaagaagccgtcttagataattattatagcaaacagaa 362  
QY 732 ACTCACTGTGGGGCTGACAGATGTTTGTCTGTGGGAAAGTACATTTGAGCATTTACTTTAGAGA 791  
Db 363 aattaccgtgggttgacggttcttgctgctggaagatacatattgagcattacttggaggga 422  
QY 792 CTTTCTGGAGTCTGTGACATGTACTTCACTGTTGGCCATCGGGTCAATTTTACCTCAT 851  
Db 423 gttcttaataatgcaaaatacattcttcatggttggtggccaaagtcattcttttaccatcat 482  
QY 852 GATAGACGACACTCTCCGGATGCTGTCTGACACCTTGAACCCCTCTTACATTTCTTACAAGT 911  
Db 483 ggtgagatatactctccagagatgctctttgataagcttggtctctctgctctctttaaagt 542  
QY 912 CTTTGTAGATCAGGTCTGAGAGAGGTGGCAGGATATCAGCATGATCGGATGAAGACCAT 971  
Db 543 gtttgagatacaagtcaggaagaggttggaagacatacagcatgatacgatgaagaccat 602  
QY 972 TGGGGAGCACATCTTGGCCACATCCAGCAGCAGGTTCGACTTCTTCTTGTGCGATGGACGT 1031  
Db 603 cggggagacatactctggccacatccagcaaggttggacttctctctctgcatggagct 662  
QY 1032 GGATCAAGTCTTTCAAGACAACCTTTCGGGTGGAAAACCTCTGGGCCAGCTGTAGCAGAGCT 1091  
Db 663 ggtacaggtcttccaaaacaaactttgggtggagacaccttggccagctcgttggtcagct 722  
QY 1092 CGAGGCTCGGTGTCAAGGCCAGTCCCAGAGAGTTTCACTATGAGAGCGGGGAACTGTCT 1151  
Db 723 acaggcctggtgtacaagggcacatctctgacgagttcactacagagggcgggaagagtc 782  
QY 1152 GCGCGGTACATCTCCAGAGGGGGATTTTACTACCGCGGCCATTTTTCGGAGG 1211  
Db 783 cgcagctacattccgtttggccagggggtattttattaccacgagccatttttggggg 842  
QY 1212 AACGCTACTCATTTCTCAACCTCACAGGGAGTGTCTTTAAGGGATCTCTCCAGGACAA 1271  
Db 843 aaacccactcaggtttctaaacatcactcaggaagtgcttcaagggaatcctccaggacaa 902  
QY 1272 GAAACATGATAGAGGCCAGTGGCATATGAGAGCCACCTCAACAAATATCTTCTCTTTT 1331  
Db 903 ggaataatgacatagaagcaggtggcatgatgaagccatctaaacaagtatttctctct 962  
QY 1332 CAACAAACCCACTAAATCTTATCTCCAGAGTATTCCTGGGACTATCAGATAGCGCTGCC 1391  
Db 963 caacaaacccactaaataattcttatcccccagaatactctgctgggaatacatatagggcatgctc 1022  
QY 1392 TTTCAAGTATTTAAAGTGTCAAGGTAGCTTTGGCAGACAAAGAGTATAATTTGGTTAGAAA 1451  
Db 1023 tgtgatatattagattgtcaagatagatttggcagaaaaaagagataaatttggttagaaa 1082  
QY 1452 TAATGTCTGA 1461  
Db 1083 taacatctga 1092

RESULT 11  
AAT04522  
ID AAT04522 standard; cDNA; 1131 BP.  
XX  
AC AAT04522;  
XX  
DT 01-MAY-1996 (first entry)  
DE  
XX Marmoset alpha-1,3-galactosyltransferase coding region.  
KW Marmoset; alpha-1,3-galactosyltransferase; Immune response; glycoprotein;  
KW alpha-galactosyl epitope; cell membrane; virus; phagocytosis; tumour;  
KW antigen processing; leukemia; lymphoma; myeloma; melanoma; carcinoma;  
KW sarcoma; vaccine; opsonisation; glycoprotein; antibody; anti-Gal; ss.



PR 15-MAY-2000; 2000US-204148P.  
PR 13-JUN-2000; 2000US-0593316.  
XX (GERO-) GERON CORP.

XX Denning C, Clark J;  
PI WPI; 2002-089848/12.  
XX P-PSDB; AAE16319.  
DR

XX New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal  
PT determinants, for xenotransplantation, and in the treatment of the  
PT human body by surgery or therapy -  
XX  
PS Claim 18; Page 63-64; 86pp; English.  
XX

CC The patent discloses immunologically compatible animal tissue, suitable  
CC for xenotransplantation into human patients. The invention also relates  
CC to an ovine tissue devoid of antibody-detectable Gal-alpha (1,3)Gal (GAL)  
CC determinants which are made by alpha(1,3) galactosyltransferase (GT). The  
CC ovine tissue is useful for treatment of human body by surgery or therapy  
CC and in xenotransplantation, by transplanting the ovine tissue into a  
CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.  
CC Polynucleotide constructs of the invention are useful for inactivating  
CC an alpha1,3 GT gene in an ovine cell. Alpha (1,3) GT sequences are useful  
CC for drug screening and for the production of GAL containing synthetic  
CC oligosaccharides. Sequences of the invention are also useful in gene  
CC therapy. The present sequence is a DNA encoding marmoset alpha (1,3)  
CC GT protein.  
XX

SQ Sequence 1131 BP; 345 A; 223 C; 283 G; 280 T; 0 other;

Query Match 46.3%; Score 693.8; DB 24; Length 1131;  
Best Local Similarity 78.1%; Pred. No. 3e-205;  
Matches 883; Conservative 0; Mismatches 197; Indels 51; Gaps 2;

QY 382 ATGAATGTCAGGGAAAGTAATCTCTGTGATGCTGATGTCATCAACCGTGGTTCGTG 441  
DB 1 atgaatgcaaaaggaaagtaattctgtcgatgctggtgtctcaactgtgattgtgtg 60  
QY 442 TTTTGGGAATATCTCAACAGA-----ATT 465  
DB 61 ttttgggaatatcatcaacagccagagagctcttcttctgtgatatcatcaactcaaaagac 120  
QY 466 CCAGAGGTTGCTGAGAACAGATGCCAGAGGACTGGTGGTTCACAGCTGGTTTAAAAAT 525  
DB 121 ccagaagttgatgaacagcagtgctcagaagagactggtggttctctggtggtttaacaat 180  
QY 526 GGGACCCACAGTTATCAAGAAGACACAGTAGAAGGACGGAGAGAAAGGGTAG----- 578  
DB 181 gggatccacaattatcaacaaggaggagagacacagacagacaaagaaaggagagaggag 240  
QY 579 -----AATCGGAGATCGCATTTGAAGACCTTCACCTATGGGACTGGTTCAATCCAAG 630  
DB 241 gaacaaaaaaggagatgacacacagagctcggtctatggagactggtttaatccaaag 300  
QY 631 AACCCGCCGATGTTTACAGAGTGACCCCTGGAAGCGCGCGATTTGTTGGGAAGGCACCT 690  
DB 301 aaacgccagaggttatgacagtgacccaatggaagggcgggttctgtg999gaagcact 360  
QY 691 TATGACACAGCTCTGCTGGAAAAGTACTACGCCACACAGAAACTCACTGTGGGGCTGACA 750  
DB 361 tacaacaagccatcctcctagaaaattattatgccaacagaaaattaccgtg999gttgacg 420  
QY 751 GTGTTTCTGCTGGGAAGTACATTGAGCATTACTTACAGAGACTTTCTGGAGTCTGCTGAC 810  
DB 421 gtttttgctattggaagatatattgagcattacttctgagaggagttcgttaacatctgcta 480  
QY 811 ATCTACTTCTATGCTGGCCATCGGTCATATTTTACGTCTATGATAGACGACACCTCCCGG 870  
DB 481 aggtacttcatggcggccacaaagtcatttcttctgtcatggcgtgatgctctcaag 540

QY 871 ATGCCTGTCTGCACACCTGAACCCCTCTACATTCCTTCTAAGTCTTTGAGATCAGGTCGTGAG 930  
DB 541 ggcgcgtttatagagcgtgggtctctcgcttctctcaaaagtgtttgaggtcaagcagag 600  
QY 931 AAGAGGTGGCAGGATATCAGCATGATGCGCATCAAGACCATTGGGGAGCACATCTCTGGCC 990  
DB 601 aagaggtggcaagacatcagca tga tgcgtatgaagacccatcggggagcaatcttggcc 660  
QY 991 CACATCAGCACGAGGTGCAGTCT 1050  
DB 661 cacatccaacagaggttgaacttctctctctctctctctctctctctctctctctctctct 720  
QY 1051 AACTTCGGGTGAAACTCTGGGCCAGCTGGTAGCAGACAGCTCCAGCGCTGGTGGTACAAG 1110  
DB 721 cattttggggtagagaccctggcgagtcggtggtctcagctacagggcgtggtgtacaag 780  
QY 1111 GCCAGTCCCGAGAAAGTTCACCTATGAGAGCGGGAACCTGTGCGCCGCTACATTCATTC 1170  
DB 781 gcagatcctctgatgaacttaccctatgagagcggaagagtcggcagcatatattccattt 840  
QY 1171 GGAGAGGGGGATTTTACTACCAACGCGGCCATTTTGGAGGAACGGCTACTCATTCTC 1230  
DB 841 gcccaggggaggtttttattaccatgcagccatttttggaggaacaccgattcaggttctc 900  
QY 1231 AACCTCACCAGGAGGTGCTTTAAGGGGATCTCCAGGACAAAGAACATGACATAGAGCC 1290  
DB 901 aacatcacccagagtgctttaagggaatctctctctggaagaagaaatgacatagaagcc 960  
QY 1291 CAGTGCATGATCAGAGCCACCTCAACAATACTTCTCTTTTCAACAACCCACTAAATC 1350  
DB 961 gagtggcatgatgaagccacccctaaagaagtattctctctcaacaacccctcaaaatc 1020  
QY 1351 CTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCTTTCAGATATTTAAAGTGTC 1410  
DB 1021 ttatctccagaataactgcgtggattatcatatagggccttcagatattaaactgtc 1080  
QY 1411 AAGGTAGCTGGCAGACAAAGAGTATAATTTGGTTAGAAATATATGTCCTGA 1461  
DB 1081 aagctatcatggcaacaaagagataatttgggttagaagaagatgctctga 1131

RESULT 13  
AAD26483  
ID AAD26483 standard; DNA; 1269 BP.  
XX  
AC AAD26483;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Pig alpha (1, 3) galactosyltransferase DNA.  
XX  
KW Pig: xenotransplantation; Gal-alpha (1,3)Gal; GAL determinant; surgery;  
KW alpha(1,3) galactosyltransferase; alpha (1, 3) GT; drug screening;  
KW gene therapy; ds.  
XX  
OS Sus scrofa.  
XX  
FH Key Location/Qualifiers  
FT CDS 16..1131  
FT /\*tag= a  
FT /product= "Pig alpha (1, 3) GT protein"  
XX  
XX WO200188096-A2.  
XX  
PD 22-NOV-2001.  
XX  
XX 14-MAY-2001; 2001WO-US15765.  
XX  
XX 15-MAY-2000; 2000US-204148P.  
PR 13-JUN-2000; 2000US-0593316.  
XX  
XX (GERO-) GERON CORP.  
XX

PI Denning C, Clark J;  
XX WPI: 2002-089848/12.  
DR P-PSDB; AAE16320.  
XX  
PT New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal  
PT determinants, for xenotransplantation, and in the treatment of the  
PT human body by surgery or therapy  
XX  
XX  
PS Claim 18; Page 65-66; 86pp; English.  
XX  
CC The patent discloses immunologically compatible animal tissue, suitable  
CC for xenotransplantation into human patients. The invention also relates  
CC to an ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal (GAL)  
CC determinants which are made by alpha(1,3) galactosyltransferase (GT). The  
CC ovine tissue is useful for treatment of human body by surgery or therapy  
CC and in xenotransplantation, by transplanting the ovine tissue into a  
CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.  
CC Polynucleotide constructs of the invention are useful for inactivating  
CC an alpha1,3 GT gene in an ovine cell. Alpha (1,3) GT sequences are useful  
CC for drug screening and for the production of GAL containing synthetic  
CC oligosaccharides. Sequences of the invention are also useful in gene  
CC therapy. The present sequence is a DNA encoding pig alpha (1,3) GT  
CC protein.  
XX  
SQ Sequence 1269 BP; 384 A; 259 C; 306 G; 320 T; 0 other;

Query Match 45.68; Score 684.2; DB 24; Length 1269;  
Best Local Similarity 76.68; Pred. No. 3.1e-202;  
Matches 880; Conservative 0; Mismatches 233; Indels 36; Gaps 2;

Qy 371 AGGAGAAATTAATGAATGCTCAAGGAAAGTAATCCCTGTTGATGCTGATTTCTCAACCG 430  
Db 5 agagaaaaataatgaatgctcaagagagtggtctgctcaatgctgctgctcaactg 64  
Qy 431 TGGTTGCTGTTTGGGAATATGTCACAGAAATTCACAGAGTT----- 474  
Db 65 taatggtgtgtgttttgggaatacatcaacagccagaggttcttctgttgataacc 124  
Qy 475 -----GGTCAGAGACAGATGGCAGAGGACTGTGTTCCCAAGCTGGT 517  
Db 125 agtcaaaaaaccagaagtttggcagcagtgctcagagagggctggttcccgagctggt 184  
Qy 518 TTAATAATGGGACCCACAGTTTATCAAGAGACACACGTAGAGAGACGGAGAGAAAAGG--- 574  
Db 185 ttaacaaatgggactcacagttaccacgaagaagacgctataggcaacgaagaagaac 244  
Qy 575 GTAGAAATGGAGATCGCATTTCAAGAGCCTCAGCTATGGGACTGGTTCAATCCAAGAAC 634  
Db 245 aaagaaagaagaacaacagaggagagcttcgctagtggactggttaatcctcagaaac 304  
Qy 635 GCCCGGATGTTTTCACAGTGAACCCGTTGGAAGCGCCGATTTGTGTGGAGGCACTTATG 694  
Db 305 gccagaggctgtgaccataaccagaaggagaggtccagtggtatgggaaggcacttaca 364  
Qy 695 ACACAGCTCTGCTGCAAAAGTACTACGCCACACAGAACTCACTGTGGGGCTGACAGTGT 754  
Db 365 acagagcgtcttagataattattatgccaaacagaaaattaccgtgggcttgagcgttt 424  
Qy 755 TTGCTGTGGGAAAGTACATTCAGGATTTACTTTAGAGACTTTTCTGGAGTCTGCTCACATGT 814  
Db 425 ttgctgtcggaagatacatatgagcattacttggaggagttcttaataatctgcaaatat 484  
Qy 815 ACTTCATGTTGGGCATCGGGTCAATTTTACGTATGATAGACGACACCTCCCGGATGC 874  
Db 485 acttcaatggttggccaaagtcattcttttacaacaagggttgggaatcctcccgagatgc 544  
Qy 875 GTGCTGTCACCTGCAACCCCTCTACATTCCTTTACAAGTCTTTGAGATCAGGCTCTGAGAAGA 934  
Db 545 ctttgatagagctgggtcctcctgcttctttaaagtggtttgagatcaagctccgagaaga 604  
Qy 935 GGTGGCAGGATATCAGCATGATGCGCATGAGACCATTTGGGACACATCTCTGCCCCACA 994

Db 605 ggtggcaagacatcagcatgatgcgcatgaagacccatcgggagcacatcctcgccaca 664  
Qy 995 TCCAGCAGGAGTGGAGTTCCTCTTCATGATGGAGTGGATCAAGTCTTTCAAGACAAC 1054  
Db 665 tccagcagaggtggagcttctctctcgcagcggagcgcagcagcagcagcagcagc 724  
Qy 1055 TCGGGGTGGAACCTCTGGGCCAGCTGTGTAGCACAGCTCCAGGCCCTGTGTGTACAGGCCA 1114  
Db 725 ttggggtggagacccctggccagtcggtggtcagctacagggcctgggtggacaggcac 784  
Qy 1115 GTCCCGAGAGTTCACCTATGAGAGCGGGGAACCTGTTCGGCCGCTACATTCATTCGGAG 1174  
Db 785 atcctgacgagttcacctacgagagggagagcgcagcagcagcagcagcagcagcagc 844  
Qy 1175 AGGGGGATTTTACTACACCGCGCCATTTTGGAGGAACGCCCTACTCACATTTCTCAACC 1234  
Db 845 aggggggatttttattaccacgcagccatttttgggggaacccacactcaggttcttaaca 904  
Qy 1235 TCACCGAGGAGTGTCTTTAAGGGATCTCTCCAGGACAGAAACATGACATAGAGCCCACT 1294  
Db 905 tctctcagagtgcttcaagggaatcctccaggaaggaagaaatgacatagagcagcaggt 964  
Qy 1295 GGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCACTAAAATCTCTAT 1354  
Db 965 ggcagtgagaaagccatctaaacaagttatttcttctcaacaacccactaaaatcttat 1024  
Qy 1355 CTCCAGAGTATTGCTGGGACTATCAGATAGAGCTTCAGATATTAAAAAGTGTCAAGG 1414  
Db 1025 cccagaaatcgtcgtggattcatatagggcagtcgtgtggatattagggattgtcaga 1084  
Qy 1415 TAGCTTGGCAGACAAAAGAGATATAATTTGGTTAGAAATAATGCTGTGACTTCAAAATGTGA 1474  
Db 1085 tagcttggcagaaaaagagataatttgggttagaataaacatctgacttttaattgtgc 1144  
Qy 1475 TGGAAACTT 1483  
Db 1145 cagcagttt 1153

RESULT 14  
AAT02892  
ID AAT02892 standard; cDNA; 1269 BP.  
XX  
AC AAT02892;  
XX  
DT 04-JUN-1996 (first entry)  
XX  
DE Porcine alpha (1,3) galactosyltransferase cDNA.  
XX  
KW Transgenic; swine; porcine; alpha (1,3) galactosyltransferase;  
KW antisense; ribozyme; Gal-alpha-1,3-Gal-beta-1-4GlcNAc; epitope;  
KW terminal; xenogenic; transplacental; rejection; gene therapy; pig; ss.  
XX  
OS Sus scrofa.  
XX  
FH Key Location/Qualifiers  
FT CDS 16..1131  
FT /\*Lag= a  
XX  
PN W09528412-A1.  
XX  
XX 26-OCT-1995.  
XX  
XX 31-MAR-1995; 95WO-US03940.  
XX  
XX 13-APR-1994; 94US-0228933.  
XX  
XX (BIOT-) BIOTRANSPLANT INC.  
XX (GEO) GEN HOSPITAL CORP.  
XX (CHIL-) INST CHILD HEALTH.  
PI Baetscher MW, Gustafsson KT, Sachs DH;

XX WPI: 1995-373759/48.  
DR P-PSDB; AAR85082.  
XX  
PT Novel transgenic alpha (1,3) galactosyl:transferase negative swine  
PT - used to produce rejection resistant cells for xenogenic  
PT transplantation  
XX  
PS Claim 11; Pages 35-37; 56pp; English.  
XX  
CC Transgenic swine in which the normal expression of the alpha (1,3)  
CC galactosyltransferase (AGT) AAR85082 is prevented, are prep'd. by  
CC inhibiting the expression of the AGT gene AAR2892 using antisense  
CC oligonucleotides or ribozyme inactivators in a pluripotent porcine  
CC embryonic stem cell. It is then inserted into a porcine oocyte  
CC (from which the pronuclear material has been removed), which is  
CC itself grown to produce the transgenic swine. Swine which do not  
CC express AGT will not produce carbohydrate moieties contg. the  
CC distinctive terminal Gal-alpha-1,3-Gal-beta-1-4GlcNAc epitope,  
CC which is a significant factor in xenogenic (esp. human) transplant  
CC rejection of swine grafts. Therefore the swine cells produced in  
CC the AGT negative transgenic swine are xenogenic transplant  
CC rejection resistant, and can therefore be used by a transplant  
CC recipient, or to provide gene therapy.  
XX  
SQ Sequence 1269 BP; 384 A; 261 C; 304 G; 320 T; 0 other;

Query Match 45.4%; Score 681; DB 16; Length 1269;  
Best Local Similarity 76.4%; Pred. No. 3.1e-201;  
Matches 878; Conservative 0; Mismatches 235; Indels 36; Gaps 2;

QY 371 AGGAGAAATTAATCAATGTCAGGAGAAAGTAATCTGTGTGATGCTGATTGTCTCAACCG 430  
DB 5 agagaaaaataagaattcaagaagaagtggttctgtaagtcgtgtgtctcaactg 64  
QY 431 TGGTTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTT----- 474  
DB 65 taatgggtgtgtttgggaatacatcaacagccagagaaggttcttctgttgatatacc 124  
QY 475 -----GGTGACAACAGATGGCAGAGGACTGTGTTCACCAAGCTGGT 517  
DB 125 agtcaaaaaaccagaagtttggcagcagtgctcagagggggtggttcccgagctggt 184  
QY 518 TTAATAATGGGACCCACAGTTATCAAGAACACACGTAGAGGAGCGAGAGAAAGG--- 574  
DB 185 ttaacaatgggactcacagttaccacgaagaagaagcgtataggcaacgaagaagaac 244  
QY 575 GTAGAAATGAGATCGGATTGAAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAGAAC 634  
DB 245 aaagaaaagaagaacaacagagagagcttcgcgtagtggactggtttaatcctcgagaaac 304  
QY 635 GCCCGGATGTTTGAACAGTCACCCGCTGGGAAGCGCGGATGTGTGGGAAGGCACTTATG 694  
DB 305 gcccaagaggtcgtagcaataaacagatggaaggctccagtcggtatgggaaggcactaca 364  
QY 695 ACACAGCTGCTCGGAAAGTACTACGCCACACAGAAACTCACTGTGGGCTGCACAGTGT 754  
DB 365 acagagcgctcttagataaattattagcaaacagaaaaattaccgtgcgcttgacggttt 424  
QY 755 TTGCTGTGGGAAGATACATTGAGCATTACTTGAAGACATTTCTTGGAGCTGTGCTGACATGT 814  
DB 425 ttgctgtcggaagatacattgagcattacttggaggagttcttaatactgcgaataacat 484  
QY 815 ACTTCATGTTGGCGCTGATATTTTACGTGATGATAGACACACCTCCCGGATGC 874  
DB 485 acttcattggtggcccaagaagtcatctttcacatcatggtggatgatattccagatgc 544  
QY 875 CTGCTGTGCAACCTCTACATTCCTTACAAGTCTTTTGAAGTCAAGCTCTGAGAAGA 934  
DB 545 cttttagatagctggttcctctcggttccctttaaagtgtttgagatcaagtccgagaaga 604  
QY 935 GGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGGAGCACATCCTGGCCACA 994

DB 605 ggtaggaagacatcagcatgtagcgaagaccaacgggggagcacatccttgccccaca 664  
QY 995 TCACAGCAGCAGGTCGACTTCTCTTCATGGCATGGACGTGGATCAAGTCTTTTCAAGACAAC 1054  
DB 665 tccagcacaggttggaactctctctcgtcagtgagcgtggaggttcctccaaaacaaact 724  
QY 1055 TCGGGGTGGAAACTCTCTGGCCACGCTGGTAGCAGACGCTCCACGCGCTGGTGTACAAAGCC 1114  
DB 725 ttgggtggagacacctgggcccagtcggtgctcagctacagggcctggtgtacaaggcac 784  
QY 1115 GTCCGAGAAGTTCACTATGAGAGGGCGGAACCTGTGCGCGCGGTACATTTCCATTCCGAG 1174  
DB 785 atcctgacgagtttcaactacagagagcggaagagtcgcgagcctacattccgtttggcc 844  
QY 1175 AGGGGATTTTACTACACGCGGCCATTTTGGAGGAACGCTTACTACATTTCTCAACC 1234  
DB 845 agggggattttttattaccacgcagccatttttgggggaacaccacactcaggttctaaaca 904  
QY 1235 TCACCAAGGGAGTCTTTAAGGGGATCTCCAGGACAAAGAACATGACATAGAACCCAGT 1294  
DB 905 tcactcagagtgcttcaagggaatcctccaggaagaaatgacatagagacgagc 964  
QY 1295 GGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCACATAAATCTCAT 1354  
DB 965 ggcattgtagaagccattcaacaagatttctcttctcaacaaacccactaaatcttat 1024  
QY 1355 CTCACAGATATTGCTGGGACTATCAGATAGGCTTCGCTTCAAGATATTAAAGTCTCAAGG 1414  
DB 1025 ccccaaatcctgctgggatacatatagggcatgctctggaattattagatgtccaaga 1084  
QY 1415 TAGCTTGGCAGACAAAGAGATATAATTTGGTTAGAAAATAATGTCTGACTTCAAAATTGTA 1474  
DB 1085 tagcttggcagaaaaagagataaatttggtagaataacatctgactttaaatgtgc 1144  
QY 1475 TGGAAACTT 1483  
DB 1145 cagcagttt 1153

RESULT 15  
AAQ93077 standard; cDNA; 1410 BP.  
XX  
XX AAQ93077;  
DT 11-DEC-1995 (first entry)  
XX Plg alpha-1,3-galactosyltransferase cDNA.  
XX Alpha-1,3-galactosyltransferase; alpha-1,3-GalT; transgenic animal;  
KW plg; hyperacute rejection; xerotransplantation; donor organ;  
KW allograft rejection; Gal epitope; gene disruption;  
KW homologous recombination; knock-out; ss.  
XX Sus scrofa.  
XX  
XX Key Location/Qualifiers  
CDS 89..1204  
FT /\*tag= a  
FT exon 1..79  
FT /\*tag= b  
FT /\*label= Exon-2  
FT exon 80..168  
FT /\*tag= c  
FT /\*label= Exon-4  
FT /\*note= "the sequence corresp. to exon-3 of the mouse gene is not present in the pig alpha-1,3-GalT gene"  
FT exon 169..204  
FT /\*tag= d  
FT /\*label= Exon-5  
FT 205..266

FT FT /\*tag= e  
FT exon /label= Exon-6  
FT 267...372  
FT /\*tag= f  
FT exon /label= Exon-7  
FT 373...510  
FT /\*tag= g  
FT exon /label= Exon-8  
FT 511...1204  
FT /\*tag= h  
FT exon /label= Exon-9

W09520661-A1.

XX XX  
XX 03-AUG-1995.

XX 27-JAN-1995; 95WO-IB00088.

XX 26-JAN-1995; 95US-0188607.

PR 27-JAN-1994; 94US-0188607.

XX (BRES-) BRESATEC LTD.

PA (SVIN-) ST VINCENT'S HOSPITAL MELBOURNE LTD.

XX Crawford RJ, Dapice AJF, Pearse MJ, Rathjen PD;  
PI Robbins AJ;

XX WPI: 1995-275446/36.

DR P-PSDB; AAR76777.

XX New alpha-1,3-galactosyltransferase and leukaemia inhibitor factor  
PT - corresp. DNA and nucleic acid constructs for inactivating the  
PT transferase gene; for eliminating hyperacute region in human  
PT transplants

XX Claim 1; Fig. 4b-4f; 184pp; English.

CC cDNA encoding porcine alpha-1,3-GalT was generated from liver RNA  
CC using primers based on conserved regions of the mouse and cattle alpha-  
CC 1,3-GalT genes. Potential sites to interrupt the alpha-1,3-GalT gene  
CC (via homologous recombination) were identified in exons 4, 7, 8 and 9.  
CC Such inactivation allows the breeding of 'knock-out' animals, e.g.  
CC pigs suitable as donors of organs to overcome hyperacute rejection  
CC problems in human xenotransplantation.

XX Sequence 1410 BP; 421 A; 287 C; 344 G; 358 T; 0 other;

Query Match 45.2%; Score 677.8; DB 16; Length 1410;  
Best Local Similarity 76.2%; Pred. No. 3.3e-200;  
Matches 876; Conservative 0; Mismatches 237; Indels 36; Gaps 2;

QY 371 AGGAGAAAATCAATGTCAGGGAAGAAAGTAATCCCTGTGATGCTGATTGCTCAACCG 430

DB 78 agggagaaataatgaatgctcaaaagagagcggctctgctcaatgctgctgctcaactg 137

QY 431 TGGTGTGCTGCTTTTGGGAATATGCAACAGAAATTCAGAGGTT----- 474

DB 138 taatggtgtgttttgggaatacatcaacagccagaggttcttctggtataacc 197

QY 475 -----GGTGAGAACAGATGGCAGAGGACTGTGGTTCCCAAGCTGGT 517

DB 198 agtcaaaaacccagagaggtggcagcagtgctcagaggggctggtggttcgagctggt 257

QY 518 TTAAGAAATGGGACCCACAGTTATCAAGAGACAACGTTAGAGGACGGAGAGAAAGG--- 574

DB 258 ttaacaatgggagctcacaggttaccacagaagaagacgctataggcaacagaaggaac 317

QY 575 GTAGAAATGGAGATCGCATTTGAGAGGCTTCAGCTATGGGACTGGTTCAATCCAAAGAAC 634

DB 318 aaagaaaagaagacaacagagagagcttccgctagcggactggttttaactctgagaac 377

QY 635 GCCCGGATGTTTTGACAGTGACCCCGGTGGAAGGCCCGGATTTGTTGGGAGGCACATTATG 694

DB 378 gccagaggtcgtgaccataaccagatggaaggtctccagtggtatgggaagaccttaca 437  
QY 695 ACACAGCTCTGCTGSAAGAAAGTACTACGCCACACAGAAACTCAGCTGTGGGGTGCACAGTGT 754  
DB 438 acagagccgtctctagataaattataatcccaaacagaaaaattaccgtggtggtcggcgttt 497  
QY 755 TTGCTGTGGGAAAGTACATTGAGCAATTTACTTTAGAGACTTTTCTGGAGTCTCTGACATGT 814  
DB 498 ttgctgtcggagaatacatatgagcattacttctggagaggttcttaataatctgcaaatatcat 557  
QY 815 ACTTCATGTTGGCCATCGGGTTCATATTTTACGTATGATAGACGACACCTCCCGGATGC 874  
DB 558 acttcatggttggccacaagtcattctttacatcatcgttggtgatgatctcccaggatgc 617  
QY 875 CTGTCGTGCACCTGAACCCCTCTACATTCCTTTACAGTCTTTTGAGATCAGGTCTCAGGAAGA 934  
DB 618 ctttgatagagctgggtccctctgcttccctttaaagtgcttgagatcaatgccgagaaga 677  
QY 935 GGTGCAGAGATATCAGCATGATGCGCATGAAGACCATTTGGGGAGCAGATCTCGGCCACACA 994  
DB 678 ggtggcaagacatcagcatgatgcgatgaagaccatcggggagacacatctctggccaca 737  
QY 995 TCCAGCAGAGGTGCGACTTCTCTTCGATGAGCGTGGATCAAGTCTTTTCAAGACAACT 1054  
DB 738 tccagcacgaggtggacttctctctcttgcatggagctggatcagggtcttccaaaacaact 797  
QY 1055 TCGGGTGGGAACCTCTGGCCAGCTGTGACACAGCTCCAGGCGCTGTGGTACAGGCGCA 1114  
DB 798 ttgggggtggagacctggggccagtcgggtggtcagctcacagggcctggtgggtacaaaggcac 857  
QY 1115 GTCCCGAGAACTTCCACCTATGAGAGGGGGAACTGTGCGCGCGGTACATTTCCATTTCGGAG 1174  
DB 858 atctcgacgaggttcacactacgagagcggaagaggtccgcagcctacattccgtttgccc 917  
QY 1175 AGGGGGATTTTACTACACCGCGGCCATTTTGGAGGAAACCCCTACTCAGATTTCTCAACC 1234  
DB 918 agggggatattttattaccacgcagccatttttgggggaacacccactcaggttctaaaca 977  
QY 1235 TCACCAGGGAGTGTCTTTAAGGGGATCTCTCCAGGACAAGAAACATGACATAGAAGCCCACT 1294  
DB 978 tcactcagggagtgcttcaagggaaatctctccaggaacaaggaataagacataagagccgagt 1037  
QY 1295 GGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCCACTAAAATCTCTAT 1354  
DB 1038 ggcattgataaagccatctctaaacaagtatatttaattctcaacaacacccactaaaacttat 1097  
QY 1355 CTCAGAGATTTGCTGGGACTATCAGATAGGCCCTGCCTTTCAGATATTAAAAGTGTCAAGG 1414  
DB 1098 cccagaaatctgctgggattatcatatagggcatgctctggtggtattaggattgtcaaga 1157  
QY 1415 TAGCTTGGCAGACAAAAGAGATATAATTTGGTTAGAAAATAATGCTGCTGACTTCAAAATGTGA 1474  
DB 1158 tagcttggcagaaaaaagagataatttgggttagaataaacaatctctgaggttttaaatgtgc 1217  
QY 1475 TGGAAACTT 1483  
DB 1218 cagcagttt 1226

Search completed: May 17, 2002, 19:35:33  
Job time: 5959 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:06:52 ; Search time 25.43 seconds  
(without alignments)  
2266.272 Million cell updates/sec

Title: US-09-863-475A-4  
Perfect score: 2136  
Sequence: 1 MITMLQDLHVKNISMSRSKS.....IKSVKVAQTKYENLVRNV 394

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTEMBL17.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2117	99.1	406	11 Q9DBU1	Q9dbul mus musculus
2	1539.5	72.1	376	6 Q28855	Q28855 platyrhini
3	803	37.6	308	6 Q9BH00	Q9bh00 macaca fasc
4	772	36.1	319	11 Q9D4R9	Q9d4r9 mus musculus
5	694.5	32.5	321	6 Q9TTN2	Q9ttn2 pan troglod
6	693	32.4	274	4 Q9UQ65	Q9uq65 homo sapien
7	690.5	32.3	354	4 Q14490	Q14490 homo sapien
8	690	32.3	348	11 Q9ET32	Q9et32 rattus norv
9	689	32.3	274	4 Q9UQ69	Q9uq69 homo sapien
10	689	32.3	274	4 Q9UQ68	Q9uq68 homo sapien
11	689	32.3	274	4 Q9U1R1	Q9u1r1 homo sapien
12	689	32.3	295	4 Q15951	Q15951 homo sapien
13	689	32.3	321	6 Q9RSM2	Q9rsm2 macaca fasc
14	688	32.1	274	4 Q9UQ63	Q9uq63 homo sapien
15	686	32.1	274	4 Q14758	Q14758 homo sapien
16	685	32.1	274	6 Q9TTN5	Q9ttn5 macaca mulla
17	684	32.0	274	4 Q9UQ67	Q9uq67 homo sapien
18	684	32.0	274	6 Q9T549	Q9t549 pan troglod
19	683	32.0	274	4 Q9UQ66	Q9uq66 homo sapien

20	683	32.0	274	6	018825	018825 p histo-blo
21	682	31.9	274	6	Q9TSM3	Q9tsm3 macaca fasc
22	681	31.9	274	4	015512	015512 h alpha-3-g
23	681	31.9	274	4	Q9UQ62	Q9uq62 homo sapien
24	680	31.8	271	6	Q9TTN4	Q9ttn4 g glycoprot
25	678.5	31.8	312	6	Q9TV74	Q9tv74 pan troglod
26	678	31.7	274	4	Q9UQ64	Q9uq64 homo sapien
27	672.5	31.5	332	11	Q9BQW2	Q9bqw2 mus musculu
28	671.5	31.4	348	6	Q9TUD9	Q9tud9 pan troglod
29	671	31.4	277	6	Q9TSM1	Q9tsm1 macaca fasc
30	671	31.4	289	6	Q9TUE0	Q9tue0 macaca mulla
31	670	31.4	327	11	Q9EQW3	Q9eqw3 mus musculu
32	667	31.2	347	4	Q9UK15	Q9uk15 homo sapien
33	664.5	31.1	262	6	Q9TTN3	Q9ttn3 pan troglod
34	662	31.0	309	6	Q9T551	Q9t551 pan troglod
35	658.5	30.8	364	6	077563	077563 sus scrofa
36	637.5	29.8	347	6	Q95158	Q95158 canis famill
37	542.5	25.4	229	4	Q9C090	Q9c090 homo sapien
38	540.5	25.3	229	4	Q9C088	Q9c088 homo sapien
39	539.5	25.3	229	4	Q99479	Q99479 homo sapien
40	539.5	25.3	229	4	Q9NY02	Q9ny02 homo sapien
41	538.5	25.2	229	4	Q9BQ97	Q9bq97 homo sapien
42	537.5	25.2	229	4	Q9NY01	Q9ny01 homo sapien
43	516.5	24.2	209	6	Q9GM53	Q9gm53 gorilla gor
44	476.5	22.3	190	6	062752	062752 macaca fasc
45	476.5	22.3	190	6	062858	062858 macaca mulla

#### ALIGNMENTS

RESULT 1  
Q9DBU1 PRELIMINARY; PRT; 406 AA.  
AC Q9DBU1;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DE GLYCOPROTEIN GALACTOSYLTRANSFERASE ALPHA 1, 3.  
GN GGTAL  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=LUNG;  
RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki L.,  
RA Hayashizaki Y.;  
RT \*Functional annotation of a full-length mouse cDNA collection.\*;  
RL Nature 409:685-690(2001).  
DR EMBL; AK004753; BAB23534.1; .  
DR MGD; MGI:95704; Ggtal  
SQ SEQUENCE 406 AA; 47866 MW; E08923F47DCD102A CRC64;

Query Match 99.1%, Score 2117, DB 11, Length 406;

Best Local Similarity 96.8%; Pred. No. 2.3e-176;

Matches 393; Conservative 1; Mismatches 0; Indels 12; Gaps 1;

```
Qy 1 MITLQDLHVNKISMSRSKSETSLPSSSGSGSOEKLIMNVKGVILLMLLIVSTVVVFWEYV 60
Db 1 MITLQDLHVNKISMSRSKSETSLPSSSGSGSOEKLIMNVKGVILLMLLIVSTVVVFWEYV 60
Qy 61 N-----RIPEYGENRWKDMWFFPSWFKNGTHSYQEDNVGGRNGRNGDRIEE 108
Db 61 NSPEGSFLWIYHTKIPEYGENRWKDMWFFPSWFKNGTHSYQEDNVGGRNGRNGDRIEE 120
Qy 109 POLWDFNPKNRPDVLTPWPKAPIVWEGTYDTALLERKYYATQKLTGLTVFAVGKYYIEH 168
Db 121 POLWDFNPKNRPDVLTPWPKAPIVWEGTYDTALLERKYYATQKLTGLTVFAVGKYYIEH 180
Qy 169 YLEDFLESADMYFMVGHVIFVYVIMDDTSRMPVHLNPLHSLQVFEIRSEKRWQDISMR 228
Db 181 YLEDFLESADMYFMVGHVIFVYVIMDDTSRMPVHLNPLHSLQVFEIRSEKRWQDISMR 240
Qy 229 MKTIGEHILAHQHEVDFLCMDVDQVFDQNFQVETLQQLVAQLQAWMYKASPEKFTYER 288
Db 241 MKTIGEHILAHQHEVDFLCMDVDQVFDQNFQVETLQQLVAQLQAWMYKASPEKFTYER 300
Qy 289 RELSAAIYPFGEGDFYHAAIFGGTPTHILNLTRECFKGILODKKHHDEAQAQWHDHSLNK 348
Db 301 RELSAAIYPFGEGDFYHAAIFGGTPTHILNLTRECFKGILODKKHHDEAQAQWHDHSLNK 360
Qy 349 YFLNKPPTKILSPCYWDYQIGLPSDIKSKVAMQTKYENLVRNV 394
Db 361 YFLNKPPTKILSPCYWDYQIGLPSDIKSKVAMQTKYENLVRNV 406
```

## RESULT 2

```
Q28855 ID Q28855 PRELIMINARY; PRT; 376 AA.
AC Q28855;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE ALPHA 1.3 GALACTOSYLTRANSFERASE.
OS Platyrrhini (New World monkeys).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates.
OX NCBI_TaxID=9479;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94331837; PubMed=8054718;
RA Henlon T.R., Macher B.A., Anaraki F., Gallili U.;
```

```
RT "Defining the minimal size of catalytically active primate alpha 1,3
RT galactosyltransferase: structure-function studies on the recombinant
RT truncated enzyme."
RL Glycobiology 4:193-201(1994).
DR EMBL: S71333; AAB31587.2; -.
SQ SEQUENCE 376 AA; 44471 MW; 13C490C0BF2FA2DF CRC64;
```

Query Match 72.1%; Score 1539.5; DB 6; Length 376;

Best Local Similarity 72.6%; Pred. No. 4.5e-126;

Matches 273; Conservative 43; Mismatches 43; Indels 17; Gaps 2;

```
Qy 36 MNVKGKIVLLMLIVSTVVVFWEYV-----RIPEYGENRWKDMWFFPSWFKN 83
Db 1 MNVKGKIVLLMLIVSTVVVFWEYVINSPEGSFLWIYHSHKNPEVDSSAQKDMWFFGFNN 60
Qy 84 GTHSYQE-----DNVEGRREKGRNGDRIEPLQWDFNPKNRPDVLTPWPKAPIVWEGT 138
Db 61 GIHYQOEEEDTDKEKEBEQKEDQEDDTLRLWDFNPKNRPDVLTPWPKAPIVWEGT 120
Qy 139 YDTALLEKYYATQKLTGLTVFAVGKYYIEHYLEDFLESADMYFMVGHVIFVYVIMDDTSR 198
Db 121 YNKAILENYAKQITVGLTVFAIGRVIIEHYLEEFVTSANRYFMVGHVIFVYVIMDDTSR 180
```

```
Qy 199 MPVHLNPLHSLQVFEIRSEKRWQDISMRMKTIGEHILAHQHEVDFLCMDVDQVFDQ 258
Db 181 APFIELGLRSLRKFVPEVKRQWQDISMRMKTIGEHILAHQHEVDFLCMDVDQVFDQ 240
Qy 259 NFGVETLQQLVAQLQAWMYKASPEKFTYERRELSAAIYPFGEGDFYHAAIFGGTPTHIL 318
Db 241 HFGVETLQQLVAQLQAWMYKADPDFTYERRKESAAIYPFGCGDFYHAAIFGGTPIQVL 300
Qy 319 NLTRCFKGILODKKHHDEAQAQWHDHSLNKYFLNKPPTKILSPCYWDYQIGLPSDIKSK 378
Db 301 NITQCFKGILODKKHHDEAQAQWHDHSLNKYFLNKPPTKILSPCYWDYHIGLPSDIKTV 360
Qy 379 KVAMQTKYENLVRNV 394
Db 361 KLSMQTKYENLVRNV 376
```

## RESULT 3

```
Q9BH00 ID Q9BH00 PRELIMINARY; PRT; 308 AA.
AC Q9BH00;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE HYPOTHETICAL 36.5 KDA PROTEIN.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FRONTAL CORTEX;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB052526; BAB21880.1; -.
KW Hypothetical protein.
SQ SEQUENCE 308 AA; 36456 MW; D7FC82471DD38A3 CRC64;
```

Query Match 37.6%; Score 803; DB 6; Length 308;

Best Local Similarity 53.1%; Pred. No. 5.3e-62;

Matches 147; Conservative 50; Mismatches 78; Indels 2; Gaps 1;

```
Qy 105 RIEPQLWDFNPKNRPDVLTPWPKAPIVWEGTYDTALLERKYYATQKLTGLTVFAVGK 164
Db 27 QVEELRLSDWFHPRKRPDVLTPWPKAPIVWEGTYDTALLERKYYATQKLTGLTVFAVGK 86
Qy 165 YIEHYLEDLFLESADMYFMVGHVIFVYVIMDDTSRMPVHLNPLHSLQVFEIRSEKRWQDI 224
Db 87 FAEEYLRFLHLSANKHFMTGYRVIFYIMVDAFLQDLPDIQPSPLRTFFKAFEDACRWLEG 146
Qy 225 SMRMKTIGEHILAHQHEVDFLCMDVDQVFDQNFQVETLQQLVAQLQAWMYKASPEK 284
Db 147 SLVVMKSLGEHITSHIQDEVDVDFLSMAVNFQNEFVETLQQLVAQLHAWMYFRNTKNE 206
Qy 285 TYERRELSAAIYPFGEGDFYHAAIFGGTPTHILNLTRECFKGILODKKHHDEAQAQWHD 344
Db 207 PYERRPTSAASTPFGCGDFYFGSLMVGTPRNILDFTEBYLNGVHIDIKNGLNSTY--EK 264
Qy 345 HLNKYFLNKPPTKILSPCYWDYQIGLPSDIKSKVA 381
Db 265 HLNKYFLNKPPTKILSPCYWDYQIGLPSDIKSKVA 301
```

## RESULT 4

```
Q9D4R9 ID Q9D4R9 PRELIMINARY; PRT; 319 AA.
AC Q9D4R9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
```

```
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 4930568D16RIK PROTEIN.
GN 4930568D16RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK016248; BAB30163.1;
DR MGD; MGI:1923109; 4930568D16RIK.
SQ SEQUENCE 319 AA; 38038 MW; 519027DC100A3B2A CRC64;

Query Match 36.1%; Score 772; DB 11; Length 319;
Best Local Similarity 49.3%; Pred. No. 2.8e-59;
Matches 138; Conservative 50; Mismatches 90; Indels 2; Gaps 1;

QY 104 DRIEPPQLDWFNPKRPNVLTVPWKAPIWEGTYDTALLEKYATQKLTGTVFVAVG 163
DB 1 DHLEEPHLSTWFDPKRPPDVIATTCGLAPVLWEGTYNREVLEQYKRLNITIGLAVFATG 100
QY 164 KYIEHYLEDLESADYFMVGHVIFVYVMDTDSRMPVHNLPLSHLQVPEIRSEKRWQD 223
DB 101 NFSKEPLRRFKSTDKFWMGVYVIFILADSTYNLPEFELGPLRTLTWRLFEEWCQD 160
QY 224 ISMRMKTIGEHILAHIOHEVDFLFCMDVQVDFQVDFNFGVETLGLVLAQLQAWYKASPEK 283
DB 161 CNLRNMNMHSHKIIQCQVQVFNLFMAVAVNQTFKNFNVETLGSVAQLHAWYFKKPRD 220
QY 284 FTYERRELSAAYIPFGEDFYHAAIFGGPTHTLNLTRCFKILQDKKHIDIAQWHD 343
DB 221 FPEERRTKSAAFIPFGEDFYHAAIFGGPTHTLNLTRCFKILQDKKHIDIAQWHD 278
QY 344 SHLNKYFLFNKPTKILSPYECWDYQIGLPSDIKSVKVAWQ 383
DB 279 RHLNKFYFINKPARVLSPYENWDRFPKTPPEIKHIKIANK 318

RESULT 5
ID Q9TTN2 PRELIMINARY; PRT; 321 AA.
AC Q9TTN2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE ABO HISTO-BLOOD GROUP A TRANSFERASE (FRAGMENT).
GN ABO.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Diamond D.C., Szalay A.A.;
RT "ABO Phylogeny."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF100976; AAF04728.1;
DR EMBL; AF100973; AAF04728.1; JOINED.
DR EMBL; AF100974; AAF04728.1; JOINED.
DR EMBL; AF100975; AAF04728.1; JOINED.
KW Transferase.
FT NON_TER 1
SQ SEQUENCE 321 AA; 37252 MW; BB8A40843885E7B7 CRC64;

Query Match 32.5%; Score 694.5; DB 6; Length 321;
Best Local Similarity 46.1%; Pred. No. 1.6e-52;
Matches 143; Conservative 46; Mismatches 108; Indels 13; Gaps 4;

QY 95 GRREKG-----RNGDRIEPPQLDWFNPKN-----RDPVLTVPWKAPIWEGTYDTAL 143
DB 12 GSLRGFCMAVREPDHJHVSPLRMVTPQPKVLTPCRKDVLTVPWLAPIWEGTFNIDI 71
QY 144 LEKYATQKLTGTVFVAVGKYIEHYLEDLESADYFMVGHVIFVYVMDTDSRMPVH 203
DB 72 LNEQFRQLNTTIGLTVAFAIKYVA-FLKFLFLEAKHPMGRVHYVYFTDQPAVPRVT 130
QY 204 LNPJLSQVPIRSEKRWQDISMMRMTTIGEHILAHIOHEVDFLFCMDVQVDFNFGVE 263
DB 131 LGTGRQLSVLEVRAYKRWQDVSMRMMEMISDFCQRRFLESDYLVCDVDMFEFRDHVGE 190
QY 264 TLGOLVAQLQAWYKASPEKFTYERRELSAAYIPFGEDFYHAAIFGGPTHTLNLTR 323
DB 191 ILTFLFTGLHPGFYSGSSREAFYERRPQSQAYIPRDEGDFYLLGGFGGSGVQEVORLTRA 250
QY 324 CFKGILODKKHIDIEAQWHDSEHLNKNYFLFNKPTKILSPYECWDYQ-IGLPSDIKSVKVAW 382
DB 251 CHQAMVQDQANGIEAVWHDSEHLNKLRLRHKPTKVLSPYELWDQQLGGLGWPSVLRKLRTA 310
QY 383 QTKENYLVNRN 392
DB 311 VPKNHQAVRN 320

RESULT 6
ID Q9UQ65 PRELIMINARY; PRT; 274 AA.
AC Q9UQ65;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE C1S-AB-SPECIFIC ALPHA 1->3 N-ACETYL GALACTOSAMINYLTRANSFERASE (FRAGMENT).
DE ABO.
GN ABO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamamoto F.;
RT "Human histo-blood group ABO gene locus alleles."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF134428; AAD26580.1;
DR EMBL; AF134427; AAD26580.1; JOINED.
KW Transferase.
FT NON_TER 1
FT NON_TER 274
SQ SEQUENCE 274 AA; 31991 MW; E78E627C028E2955 CRC64;
```



```
RESULT 9
Q9U069 PRELIMINARY; PRT; 274 AA.
AC Q9U069:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE AI-SPECIFIC ALPHA 1->3 N-ACETYLGLACTOSAMINYLTRANSFERASE (FRAGMENT).
GN ABO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamamoto F.;
RT "Human histo-blood group ABO gene locus alleles.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF134418; AAD26575.1; -.
DR EMBL; AF134417; AAD26575.1; JOINED.
KW Transferase.
FT NON_TER 1
FT NON_TER 274
SQ SEQUENCE 274 AA; 31961 MW; 12974A4C48954AB6 CRC64;

Query Match 32.3%; Score 689; DB 4; Length 274;
Best Local Similarity 49.3%; Pred. No. 3.9e-52;
Matches 135; Conservative 43; Mismatches 94; Indels 2; Gaps 2;

QY 120 RPDVLVTWPKAPIVWEGTYDTALLEKYATOKLTGTLVFAVGKYIEHYLEDLFLESADM 179
DB 1 RKDVLVTPWLPAPIVWEGTFNIDILNEQFRLQNTTIGLTVFAIKKYVA-FLKLFLETAEK 59
QY 180 YFMVGHVRVIFYVMIDTDSRMPVYVHLNPLHSLOVFEIRSEKRWODISMRRMKTIGEHLAH 239
DB 60 HFVGHVRVHYVFTDQAAVPRVTLGTGRQLSVLEVRAVKRWODVSMRRMEMISDFCERR 119
QY 240 IQEVDVFLFCMDVDQVDFQDNFGVETLGQLVAQLQAWMYKASPEKTYERRELSAAVYIPFG 299
DB 120 FLSEVDYLCVDVDMFDRHVGVEILTPLFGTLHPGFYGSREAFYERRPQSOAYIPKD 179
QY 300 EGDFFYYHAAIFGGTPTTHILNLTRECFKGILODKKKHIDIEAQWHDHSHLNKYFLFNKPTKIL 359
DB 180 EGDFFYLLGGFFGGSVQEVORLTRACHQAMVVDQANGIEAVWHDHSHLNKYLRLRHKPTKVL 239
QY 360 SPEYCWQDYQ-IGLPSDIKSVKVAWQTKYENLVN 392
DB 240 SPEYLDQQLLGPWPAVLRLKRLFTAVPKNHQAVRN 273

RESULT 11
Q9UIR1 PRELIMINARY; PRT; 274 AA.
AC Q9UIR1:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ABO GLYCOSYLTRANSFERASE (FRAGMENT).
GN ABO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yip S.P.;
RT "Single-tube multiplex PCR-SSCP analysis distinguishes seven common
RT ABO alleles and readily identifies new alleles.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF182746; AAF07060.1; -.
DR EMBL; AF182745; AAF07060.1; JOINED.
KW Transferase.
FT NON_TER 1
FT NON_TER 274
SQ SEQUENCE 274 AA; 32020 MW; 11F74A4C48954AB6 CRC64;

Query Match 32.3%; Score 689; DB 4; Length 274;
Best Local Similarity 49.3%; Pred. No. 3.9e-52;
Matches 135; Conservative 43; Mismatches 94; Indels 2; Gaps 2;

QY 120 RPDVLVTWPKAPIVWEGTYDTALLEKYATOKLTGTLVFAVGKYIEHYLEDLFLESADM 179
DB 1 RKDVLVTPWLPAPIVWEGTFNIDILNEQFRLQNTTIGLTVFAIKKYVA-FLKLFLETAEK 59
QY 180 YFMVGHVRVIFYVMIDTDSRMPVYVHLNPLHSLOVFEIRSEKRWODISMRRMKTIGEHLAH 239
DB 60 HFVGHVRVHYVFTDQAAVPRVTLGTGRQLSVLEVRAVKRWODVSMRRMEMISDFCERR 119
QY 240 IQEVDVFLFCMDVDQVDFQDNFGVETLGQLVAQLQAWMYKASPEKTYERRELSAAVYIPFG 299
DB 120 FLSEVDYLCVDVDMFDRHVGVEILTPLFGTLHPGFYGSREAFYERRPQSOAYIPKD 179
```

```
FT NON_TER 1
FT NON_TER 274
SQ SEQUENCE 274 AA; 31977 MW; 71BE047A648E2D8A CRC64;

Query Match 32.3%; Score 689; DB 4; Length 274;
Best Local Similarity 49.3%; Pred. No. 3.9e-52;
Matches 135; Conservative 43; Mismatches 94; Indels 2; Gaps 2;

QY 120 RPDVLVTWPKAPIVWEGTYDTALLEKYATOKLTGTLVFAVGKYIEHYLEDLFLESADM 179
DB 1 RKDVLVTPWLPAPIVWEGTFNIDILNEQFRLQNTTIGLTVFAIKKYVA-FLKLFLETAEK 59
QY 180 YFMVGHVRVIFYVMIDTDSRMPVYVHLNPLHSLOVFEIRSEKRWODISMRRMKTIGEHLAH 239
DB 60 HFVGHVRVHYVFTDQAAVPRVTLGTGRQLSVLEVRAVKRWODVSMRRMEMISDFCERR 119
QY 240 IQEVDVFLFCMDVDQVDFQDNFGVETLGQLVAQLQAWMYKASPEKTYERRELSAAVYIPFG 299
DB 120 FLSEVDYLCVDVDMFDRHVGVEILTPLFGTLHPGFYGSREAFYERRPQSOAYIPKD 179
QY 300 EGDFFYYHAAIFGGTPTTHILNLTRECFKGILODKKKHIDIEAQWHDHSHLNKYFLFNKPTKIL 359
DB 180 EGDFFYLLGGFFGGSVQEVORLTRACHQAMVVDQANGIEAVWHDHSHLNKYLRLRHKPTKVL 239
QY 360 SPEYCWQDYQ-IGLPSDIKSVKVAWQTKYENLVN 392
DB 240 SPEYLDQQLLGPWPAVLRLKRLFTAVPKNHQAVRN 273

RESULT 11
Q9UIR1 PRELIMINARY; PRT; 274 AA.
AC Q9UIR1:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ABO GLYCOSYLTRANSFERASE (FRAGMENT).
GN ABO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yip S.P.;
RT "Single-tube multiplex PCR-SSCP analysis distinguishes seven common
RT ABO alleles and readily identifies new alleles.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF182746; AAF07060.1; -.
DR EMBL; AF182745; AAF07060.1; JOINED.
KW Transferase.
FT NON_TER 1
FT NON_TER 274
SQ SEQUENCE 274 AA; 32020 MW; 11F74A4C48954AB6 CRC64;

Query Match 32.3%; Score 689; DB 4; Length 274;
Best Local Similarity 49.3%; Pred. No. 3.9e-52;
Matches 135; Conservative 43; Mismatches 94; Indels 2; Gaps 2;

QY 120 RPDVLVTWPKAPIVWEGTYDTALLEKYATOKLTGTLVFAVGKYIEHYLEDLFLESADM 179
DB 1 RKDVLVTPWLPAPIVWEGTFNIDILNEQFRLQNTTIGLTVFAIKKYVA-FLKLFLETAEK 59
QY 180 YFMVGHVRVIFYVMIDTDSRMPVYVHLNPLHSLOVFEIRSEKRWODISMRRMKTIGEHLAH 239
DB 60 HFVGHVRVHYVFTDQAAVPRVTLGTGRQLSVLEVRAVKRWODVSMRRMEMISDFCERR 119
QY 240 IQEVDVFLFCMDVDQVDFQDNFGVETLGQLVAQLQAWMYKASPEKTYERRELSAAVYIPFG 299
DB 120 FLSEVDYLCVDVDMFDRHVGVEILTPLFGTLHPGFYGSREAFYERRPQSOAYIPKD 179
```









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:05:07 ; Search time 15.93 seconds  
(without alignments)  
1884.041 Million cell updates/sec

Title: US-09-863-475A-4  
Perfect score: 2136  
Sequence: 1 MITMLQDLHVNKISMSRSKS.....IKSVKVAWQTKYENLVNRNV 394  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues  
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2136	100.0	394	2 A34417	alpha-1,3-mannosyl
2	1950	91.3	371	2 I49698	alpha-1,3-galactos
3	1533.5	71.8	376	2 A56480	N-acetylglactosamin
4	1510	70.7	371	2 I46583	alpha-1,3-galactos
5	1476.5	69.1	368	2 A44785	N-acetylglactosamin
6	689	32.3	295	2 PC1120	glucosylgalactose a
7	688	32.2	353	1 A34933	glucosylgalactose a
8	637.5	29.8	347	2 JC6126	glucoside alpha-N-
9	477.5	22.4	189	2 PC1168	histo-blood group
10	476.5	22.3	189	2 PC1166	histo-blood group
11	469.5	22.0	189	2 PC1165	glucosylgalactose a
12	467.5	21.9	189	2 PC1173	histo-blood group
13	464.5	21.7	189	2 PC1171	histo-blood group
14	462.5	21.7	189	2 PC1169	histo-blood group
15	462.5	21.7	189	2 PC1170	histo-blood group
16	462.5	21.7	189	2 PC1172	histo-blood group
17	114	5.3	26	2 A39769	N-acetylglactosamin
18	99.5	4.7	301	2 G85752	outer membrane pro
19	98	4.6	481	2 T27435	hypothetical prote
20	97	4.5	569	2 A43317	germ cell-less pro
21	96.5	4.5	301	2 B64881	probable outer mem
22	96	4.5	455	2 T29555	hypothetical prote
23	95.5	4.5	361	2 G64354	hypothetical prote
24	95	4.4	505	2 A32261	agarase (EC 3.2.1.
25	94.5	4.4	1078	2 T30879	dynlein heavy chain
26	93	4.4	244	2 H81905	probable phosphoad
27	91.5	4.3	389	2 B84462	hypothetical prote
28	91.5	4.3	455	2 D75043	serine--trRNA synthet
29	91.5	4.3	460	2 G71117	serine--trRNA ligas

30	90.5	4.2	496	2 A69691	ATP-dependent DNA
31	90	4.2	385	2 A81158	cystathionine beta
32	90	4.2	1478	2 T38712	ABC transporter Sp
33	89.5	4.2	612	2 T18463	hypothetical prote
34	89	4.2	895	2 F75608	conserved hypotheat
35	88.5	4.1	460	2 T48061	alpha galactosyltr
36	88.5	4.1	2896	2 T30939	hemocyanin G-type
37	88	4.1	393	2 S61659	KTR1 protein - yea
38	88	4.1	1066	2 T30903	arachidonate 8-lip
39	87.5	4.1	354	2 E81281	hypothetical prote
40	87.5	4.1	1941	2 T30554	ubiquitin-protein
41	87	4.1	385	2 B81949	cystathionine beta
42	87	4.1	865	2 D72206	valine--trRNA ligas
43	87	4.1	1102	2 S55100	hypothetical prote
44	87	4.1	1169	2 T30207	dynlein heavy chain
45	86.5	4.0	633	2 F75107	abc transporter, A

ALIGNMENTS

RESULT 1

A34417

alpha-1,3-mannosyl-glycoprotein beta-1,4-N-acetylglucosaminyltransferase (EC 2.4.1.14  
C:Species: Mus musculus (house mouse)  
C:Date: 15-Jun-1990 #sequence\_revision 15-Jun-1990 #text\_change 29-Sep-1999  
C:Accession: A34417  
R:Larsen, R.D.; Rajan, V.P.; Ruff, M.M.; Kukowska-Latallo, J.; Cummings, R.D.; Lowe,  
Proc. Natl. Acad. Sci. U.S.A. 86, 8227-8231, 1989  
A:Title: Isolation of a cDNA encoding a murine UDPgalactose:beta-D-galactosyl-1,4-N-a  
A:Reference number: A34417; MUID:90046769  
A:Accession: A34417  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-394 <LAR>

A:Cross-references: GB:M26925; NID:g193419; PIDN:AAA37657.1; PID:g309242  
C:Superfamily: histo-blood group 1 transferase  
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 100.0%; Score 2136; DB 2; Length 394;  
Best Local Similarity 100.0%; Pred. No. 2.7e-174;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MITMLQDLHVNKISMSRSKSETSLPSSRSQSEKIMNVKGVILLMLIVSTVVVWFVEYV	60
Db	1	MITMLQDLHVNKISMSRSKSETSLPSSRSQSEKIMNVKGVILLMLIVSTVVVWFVEYV	60
Qy	61	NRPEVGENRWQKDWPFPSFKNGTHSYQEDNVEGRREKGRNGRIEPPQLMDWPNPKNR	120
Db	61	NRPEVGENRWQKDWPFPSFKNGTHSYQEDNVEGRREKGRNGRIEPPQLMDWPNPKNR	120
Qy	121	PDVLTVPWKAPIVWEGTYDTALLEKYYATOKLTGLTVFAVGKIEHYLEDLFESADMY	180
Db	121	PDVLTVPWKAPIVWEGTYDTALLEKYYATOKLTGLTVFAVGKIEHYLEDLFESADMY	180
Qy	181	FMVGHRTVFFYVIMDDTSMPVHNLPLHSLOVFEIRSEKRWQDSIMMRMKTIGEHILAH	240
Db	181	FMVGHRTVFFYVIMDDTSMPVHNLPLHSLOVFEIRSEKRWQDSIMMRMKTIGEHILAH	240
Qy	241	QHEVDFLFCMDVDQVQDNFGVETLGQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGE	300
Db	241	QHEVDFLFCMDVDQVQDNFGVETLGQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGE	300
Qy	301	GDFYFHAHAFGGTPTHILNLTRECFKGLQDKKHIDIAQWIDESHLNKYFLFNKPTKILS	360
Db	301	GDFYFHAHAFGGTPTHILNLTRECFKGLQDKKHIDIAQWIDESHLNKYFLFNKPTKILS	360
Qy	361	PEYCWQDYOIGLPSDIKSVKVAWQTKYENLVNRNV	394
Db	361	PEYCWQDYOIGLPSDIKSVKVAWQTKYENLVNRNV	394

RESULT 2  
149698  
alpha-1,3-galactosyltransferase - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 29-Sep-1999  
C:Accession: I49698  
R:Joziasse, D.H.; Shaper, N.L.; Kim, D.; van der Eijnden, D.H.; Shaper, J.H.  
J. Biol. Chem. 267, 5534-5541, 1992  
A:Title: Murine alpha-1,3-galactosyltransferase: A single gene locus specifies four isoforms  
A:Reference number: A42117; MUID:92184813  
A:Accession: I49698  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-371 <RES>  
A:Cross-references: GB:M85153; NID:g193563; PIDN:AAA37711.1; PID:g457142  
C:Genetics:  
A:Gene: Gcta-1  
C:Superfamily: histo-blood group 1 transferase

Query Match 91.3%; Score 1950; DB 2; Length 371;  
Best Local Similarity 96.5%; Pred. No. 1.7e-158;  
Matches 358; Conservative 1; Mismatches 0; Indels 12; Gaps 1;  
Qy 36 MNVKGKVLILLMLIVSTVVVWFVEYVN-----RIPEVGENRWQKOWMFPSPFN 83  
Db 1 MNVKGKVLILLMLIVSTVVVWFVEYVNPDGSLFWIYHTRKIPEVGENRWQKOWMFPSPFN 60  
Qy 84 GTHSYQEDNVEGRREKGRNGRIEIPQWDFNPNKRPDVLTVTPWKAPIVWEGTYD 143  
Db 61 GTHSYQEDNVEGRREKGRNGRIEIPQWDFNPNKRPDVLTVTPWKAPIVWEGTYD 120  
Qy 144 LEKYATQKLTGLTVFAVGKYEHLDEFLSADMYFMVGHVRVIFVYVIMDDTSRMPV 203  
Db 121 LEKYATQKLTGLTVFAVGKYEHLDEFLSADMYFMVGHVRVIFVYVIMDDTSRMPV 180  
Qy 204 LNPLHSLOVFEIRSEKRWQDSMMRMKTIGEHILAHIOHEVDVDFLCMDVQVDFN 263  
Db 181 LNPLHSLOVFEIRSEKRWQDSMMRMKTIGEHILAHIOHEVDVDFLCMDVQVDFN 240  
Qy 264 TLGOLVAQLQAWKYKASPEKTYERRELSSAAYIPFGGDFYHYHAAIFGGTPTTH 323  
Db 241 TLGOLVAQLQAWKYKASPEKTYERRELSSAAYIPFGGDFYHYHAAIFGGTPTTH 300  
Qy 324 CFKGLQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILSPYCWYDYOIGLPSDIKSV 383  
Db 301 CFKGLQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILSPYCWYDYOIGLPSDIKSV 360  
Qy 384 TKEYNLVRNV 394  
Db 361 TKEYNLVRNV 371

RESULT 3  
A56480  
N-acetylglucosaminide alpha-1,3-galactosyltransferase (EC 2.4.1.151) - marmoset  
C:Species: Callithrix sp.  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 29-Sep-1999  
C:Accession: A56480  
R:Henion, T.R.; Macher, B.A.; Anaraki, F.; Galili, U.  
Glycobiology 4, 193-201, 1994  
A:Title: Defining the minimal size of catalytically active primate alpha1,3 galactosyltransferase  
A:Reference number: A56480; MUID:94331837  
A:Accession: A56480  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-376 <HEN>  
A:Cross-references: GB:S71333; NID:g558051; PIDN:AAB31587.1; PID:g558052  
A:Note: authors translated the codon GTG for residue 251 as Ser  
C:Superfamily: histo-blood group 1 transferase  
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 71.8%; Score 1533.5; DB 2; Length 376;  
Best Local Similarity 72.3%; Pred. No. 5.5e-123;  
Matches 272; Conservative 43; Mismatches 44; Indels 17; Gaps 2;  
Qy 36 MNVKGKVLILLMLIVSTVVVWFVEYVN-----RIPEVGENRWQKOWMFPSPFN 83  
Db 1 MNVKGKVLILLMLIVSTVVVWFVEYVINSPEGSLFWIYHSHKNPEYDSSAOKDMWFGMFWN 60  
Qy 84 GTHSYQEDNVEGRREKGRNGRIEIPQWDFNPNKRPDVLTVTPWKAPIVWEGT 138  
Db 61 GTHSYQEDNVEGRREKGRNGRIEIPQWDFNPNKRPDVLTVTPWKAPIVWEGT 120  
Qy 139 YDTALLEKYATQKLTGLTVFAVGKYEHLDEFLSADMYFMVGHVRVIFVYVIMDDTSR 198  
Db 121 YKAILNENYAKOKITVGLTVFAIGRYIEHYLEEFVTSANRYFMVGHVKVIFVYVIMDDVSK 180  
Qy 199 MPVHLNPLHSLOVFEIRSEKRWQDSMMRMKTIGEHILAHIOHEVDVDFLCMDVQVDF 258  
Db 181 APFIELGLPLRSFKVEFKPEKRWQDSMMRMKTIGEHILAHIOHEVDVDFLCMDVQVDF 240  
Qy 259 NFGVETLQOLVAQLQAWKYKASPEKTYERRELSSAAYIPFGGDFYHYHAAIFGGTPTTH 318  
Db 241 HFGVETLQOLVAQLQAWKYKADPDQDFTYERRKESAAIPFGGDFYHYHAAIFGGTPTTH 300  
Qy 319 NUTRECFKGLQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILSPYCWYDYOIGLPSDIKSV 378  
Db 301 NITQCFKGLQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILSPYCWYDYOIGLPSDIKSV 360  
Qy 379 KVAWOTKEYNLVRNV 394  
Db 361 KUSWOTKEYNLVRNV 376

RESULT 4  
I46583  
alpha-1,3-galactosyltransferase - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 29-Sep-1999  
C:Accession: I46583  
R:Strahan, K.M.; Gu, F.; Preece, A.F.; Gustavsson, L.; Andersson, K.  
Immunogenetics 41, 101-105, 1995  
A:Title: cDNA sequence and chromosome localization of pig alpha 1,3 galactosyltransferase  
A:Reference number: I46583; MUID:95104914  
A:Accession: I46583  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-371 <STR>  
A:Cross-references: GB:L36152; NID:g642635; PIDN:AAA73558.1; PID:g642636  
C:Genetics:  
A:Gene: GGTA1  
C:Superfamily: histo-blood group 1 transferase

Query Match 70.7%; Score 1510; DB 2; Length 371;  
Best Local Similarity 72.7%; Pred. No. 5.4e-121;  
Matches 271; Conservative 48; Mismatches 38; Indels 16; Gaps 4;  
Qy 36 MNVKGKVLILLMLIVSTVVVWFVEYVN-----RIPEVGENRWQKOWMFPSPFN 83  
Db 1 MNVKGKVLILLMLIVSTVVVWFVEYVINSPEGSLFWIYQSKNPEVGSSA-QRGWMPSPFN 59  
Qy 84 GTHSY-QEDNVEGRREKGRNGRIEIPQWDFNPNKRPDVLTVTPWKAPIVWEGTYD 141  
Db 60 GTHSYHEEDDAIGNEKEQRKEDNRGELP-LVDWFPNPKRPVTVITRWKAPVWEGTYNR 118  
Qy 142 ALLEKYATQKLTGLTVFAVGKYEHLDEFLSADMYFMVGHVRVIFVYVIMDDTSRMPV 201  
Db 119 AVLDNRYAKQKLTGLTVFAVGRIEYHLEEFISANTYFMVGHVKVIFVYVIMDDISRMPL 178  
Qy 202 VHLNPLHSLOVFEIRSEKRWQDSMMRMKTIGEHILAHIOHEVDVDFLCMDVQVDFN 261  
Db 179 IELGSLRSPKVEIKSEKRWQDSMMRMKTIGEHILAHIOHEVDVDFLCMDVQVDFN 238

QY 262 VETLGQVLAQLQAWMYKASPEKFTYERRELSAAYIPFEGDFYHYHAAIFGGTPTTHILNLT 321  
Db 239 VETLGQVLAQLQAWMYKAHPDEFTYERRKESAAIYPGQGFYHYHAAIFGGTPTQVLNIT 298  
QY 322 RECFCGTLQDKKHIDIAQWHDHSHLNKPYLFNKPRTKILSPYCWQDYOIGLPSDIKSKVKA 381  
Db 299 OECFCGTLQDKKHIDIAQWHDHSHLNKPYLFNKPRTKILSPYCWQDHYGHGMSVDIRIVKIA 358  
QY 382 WOTKEYNLVRNV 394  
Db 359 WOKKEYNLVRNI 371

RESULT 5  
A44785  
N:acetylglucosaminyl 3-alpha-galactosyltransferase (EC 2.4.1.124) - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 19-Mar-1993 #sequence\_revision 19-Mar-1993 #text\_change 29-Sep-1999  
C:Accession: A44785  
R:Joelasse, D.H.; Shaper, J.H.; Van den Eljnden, D.H.; Van Tunen, A.J.; Shaper, N.L.  
J. Biol. Chem. 264, 14290-14297, 1989  
A:Title: Bovine alpha1->3-galactosyltransferase: Isolation and characterization of a cDN  
A:Reference number: A44785; MUID:89340543  
A:Accession: A44785  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-368 <JOZ>  
C:Superfamily: histo-blood group 1 transferase  
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 69.1%; Score 1476.5; DB 2; Length 368;  
Best Local Similarity 72.2%; Pred. No. 3.8e-118;  
Matches 268; Conservative 41; Mismatches 47; Indels 15; Gaps 3;

QY 36 MNYKGVILLMLIVSTVVVWFYEV-----NRIPVEGENRWQKDMFPSPFKN 83  
Db 1 MNYKGVILSMLVSTVIYVWFYEVIIHSPEGLFWINPNSRNPVEGGSGIQKGWMLPRWFNN 60  
QY 84 GTHSYOEDNVEGRKGRNGRIEELPOLWDFNPKRNPVDLTVTPKAPITVWECTYDTAL 143  
Db 61 GYH -BEDGDINEKEQORNEDE-SKULSDWFPFRPEVVTTKKAPVWVEGTYNRAV 117  
QY 144 LEKYATQKLTGLTVFAVGKYIEHYLEDFLESADMYFMVGRHVIYFVIMDDTSRMPVVH 203  
Db 118 LDNYAKOKITVGLTVFAVGRIYIEHYLEELPTSANKHFMVGHVPYIFVIMDDYSRMPLE 177  
QY 204 LNPFLSHQVFEIRSEKRWQDISMRMKTIGEHLAHIQHEVDFLFCMDVDQVQDNFQVGE 263  
Db 178 LGPLRSFKVFKIPKRWQDISMRMKTIGEHLAHIQHEVDFLFCMDVDQVQDNFQVGE 237  
QY 264 TLGQVLAQLQAWMYKASPEKFTYERRELSAAYIPFEGDFYHYHAAIFGGTPTTHILNLTRE 323  
Db 238 TLGESVLAQLQAWMYKADPNDFTYERRKESAAIYPFEGDFYHYHAAIFGGTPTQVLNITQE 297  
QY 324 CFFKGILODKKHIDIAQWHDHSHLNKPYLFNKPRTKILSPYCWQDYOIGLPSDIKSKVWAWQ 383  
Db 298 CFFKGILODKKHIDIAQWHDHSHLNKPYLFNKPRTKILSPYCWQDHYGHGLPADIKLVKMSWQ 357  
QY 384 TKEYNLVRNV 394  
Db 358 TKEYNLVRNV 368

RESULT 6  
PC1120  
fucosylgalactose alpha-N-acetylglucosaminyltransferase (EC 2.4.1.40) A2 allele [valid  
N:Alternate names: histo-blood group A2 transferase  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 08-Sep-2000  
C:Accession: PC1120  
R:Yamamoto, F.; McNeill, P.D.; Hakomori, S.

Biochem. Biophys. Res. Commun. 187, 366-374, 1992  
A:Title: Human histo-blood group A2 transferase coded by A2 allele, one of the A sub  
the carboxyl terminal.  
A:Reference number: PC1120; MUID:92392351  
A:Accession: PC1120  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-295 <YAM>  
A:Cross-references: GB:S44054; NID:g255066; PIDN:AAB23167.1; PID:g255067  
C:Superfamily: histo-blood group 1 transferase  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 32.3%; Score 689; DB 2; Length 295;  
Best Local Similarity 49.3%; Pred. No. 3.6e-51;  
Matches 135; Conservative 43; Mismatches 94; Indels 2; Gaps 2;  
QY 120 RPDVLTVPWKAPITVWEGTYDTALLEKYATQKLTGLTVFAVGKYIEHYLEDFLESADM 179  
Db 1 RKDVLVTPWLPAPITVWEGTENNIDILNQFRLQNTTIGLTVFAIKKYVA-FLKLFLETAEK 59  
QY 180 YFVGVHRVIFYVMIDDDTSRMPVPHLNPFLSHQVFEIRSEKRWQDISMRMKTIGEHLAH 239  
Db 60 HEMVGRHVRHYVFTDQLAAVPRVTLGTGRQLSVLEVRAYKRWQDVSMMRMEMISDFCERR 119  
QY 240 IQHEVDFLFCMDVDQVQDNFQVETLGLQVLAQLQAWMYKASPEKFTYERRELSAAYIPFG 299  
Db 120 FLSEVDVLCVDVDMEDFRDHVGVEILTFLCTLHPGFYSGSREAFYERRRQSOAYIPKD 179  
QY 300 EGFYHYHAAIFGGTPTTHILNLTRECFCGTLQDKKHIDIAQWHDHSHLNKPYLFNKPRTKIL 359  
Db 180 EGFYHYLGGFFGGSVQEVQRLTRACHQAMVMDQANGIEAVVHDESHLNKPYLLRHKPTKV 239  
QY 360 SPEYCDWYO-IGLPSDIKSKVWAWQTKYENLVRN 392  
Db 240 SPEYLDQQLGLGMPAVLRKURFTAVPRNHQAVRN 273

RESULT 7  
A34933  
fucosylgalactose alpha-N-acetylglucosaminyltransferase (EC 2.4.1.40) A1 allele [val  
N:Alternate names: alpha-3-N-acetylglucosaminyltransferase; blood-group substance A  
e: histo-blood group A glycosyltransferase  
C:Species: Homo sapiens (man)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 08-Sep-2000  
C:Accession: A34933; S09593; SI3173; PC1164  
R:Yamamoto, F.; Marken, J.; Tsuji, T.; White, T.; Clausen, H.; Hakomori, S.  
J. Biol. Chem. 265, 1146-1151, 1990  
A:Title: Cloning and characterization of DNA complementary to human UDP-GalNAc: Fucal  
A:Reference number: A34933; MUID:90110098  
A:Accession: A34933  
A:Molecule type: mRNA  
A:Residues: 1-353 <YAM1>  
A:Cross-references: GB:J05175; NID:g340077; PIDN:AAA36792.1; PID:g340078  
R:Yamamoto, F.I.; Clausen, H.; White, T.; Marken, J.; Hakomori, S.I.  
Nature 345, 229-233, 1990  
A:Title: Molecular genetic basis of the histo-blood group ABO system.  
A:Reference number: S09593; MUID:90238543  
A:Accession: S09593  
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
A:Molecule type: mRNA  
A:Residues: 1-353 <YAM2>  
R:Navaratnam, N.; Findlay, J.B.C.; Keen, J.N.; Watkins, W.M.  
Biochem. J. 271, 93-98, 1990  
A:Title: Purification, properties and partial amino acid sequence of the blood-group-A  
A:Reference number: SI3173; MUID:91024951  
A:Accession: SI3173  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 'X', 65-73, 'IS', 76-77 <NAV>  
R:Kominato, Y.; McNeill, P.D.; Yamamoto, M.; Russell, M.; Hakomori, S.; Yamamoto, F.  
Biochem. Biophys. Res. Commun. 189, 154-164, 1992  
A:Title: Animal histo-blood group ABO genes.

A:Reference number: PC1164; MUID:93080551  
A:Accession: PC1164  
A:Molecule type: DNA  
A:Residues: 145-154,'p',156-333 <KOM>  
C:Comment: This enzyme forms group A blood type determinants from H antigen determinants minor subtypes, for example A2 (see PIR:PC1120). Inactive alleles are responsible for group A subtypes.  
C:Genetics:  
A:Gene: GDB:ABO  
A:Cross-references: GDB:118956; OMIM:110300  
A:Map position: 9q34.1-9q34.2  
C:Complex: homodimer  
C:Function:  
A:Description: transfers N-acetylgalactosamine from UDP-N-acetylgalactosamine to glycoproteins  
C:Superfamily: histo-blood group I transferase  
C:Keywords: blood group; glycoprotein; glycosyltransferase; Golgi apparatus; hexosyltransferase  
F:1-11/Domain: intracellular #status predicted <INT>  
F:12-37/Domain: transmembrane #status predicted <TM>  
F:38-353/Domain: trans-Golgi network lumenal #status predicted <LUM>  
F:112/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 32.2%; Score 688; DB 1; Length 353;  
Best Local Similarity 45.6%; Pred. No. 5.6e-51;  
Matches 141; Conservative 48; Mismatches 108; Indels 12; Gaps 4;

Qy 95 GRREKG-----RNGRIEPPQLMDWPNKPNR-----PDVLTVPKAPIVWEGTYDTALL 144  
Db 45 GSLRGFCMAVREPDHLQVSLPRMVYPPQKVLTPWKDVLVVTPLWLPAPIVWEGTFNIDIL 104

Qy 145 EKYATOKLTGLTVFAVGKIEHYLEDLFLESADMYFMVGHVIRFYVIMDDTSRMPVHL 204  
Db 105 NEQPLQNTTGLTVFAIKYVA-FLKLFLEAEKHEFMVGHVIRFYVIMDDTSRMPVHL 163

Qy 205 NPLHSLOVFEIRSEKRWQDISMRMKTIGEHILAHIOHEVDFLCMDVDVQFQDNFGVET 264  
Db 164 GTGQLSVLEVYAKRWQDSMRMEMISDFCERRFLESDVYLCVDVDMFDRHVGVEI 223

Qy 265 LGQLVAQLQAWYKASPEKTYERRELSAAYIPFEGGDFYHAAIFGCTPHTLNLTRC 324  
Db 224 LTPFLGLHPFGYSSREAFYERRPQSOAVIPKDEGDFYVLGGFFGSGVEVQLTRAC 283

Qy 325 FKGLTQDKKHIDIAQWHDHSHLNKYLFLNPKTKLSPEYCHDYQ-IGLPSDIKSVKVAWQ 383  
Db 284 HQAMVDOANGIEAVWHDHSHLNKYLRLHKPTKVLSPYLMDQQLGMPAVLRKLRTAV 343

Qy 384 TKEYNLVRN 392  
Db 344 PKNHQAVERN 352

RESULT 8  
JC6126  
globoSide alpha-N-acetylgalactosaminyltransferase (EC 2.4.1.88) - dog  
N:Alternate names: forssman glycolipid synthetase; UDP-GalNAc:globoSide alpha-1,3-N-acetylglucosaminyltransferase (dog)  
C:Species: Canis lupus familiaris (dog)  
C:Date: 23-Mar-1997 #sequence\_revision 09-May-1997 #text\_change 29-Sep-1999  
C:Accession: JC6126  
R:Haslam, D.B.; Baenziger, J.U.  
Proc. Natl. Acad. Sci. U.S.A. 93, 10697-10702, 1996  
A:Title: Expression cloning of Forssman glycolipid synthetase: A novel member of the histo-blood group I transferase family  
A:Reference number: JC6126; MUID:9708066  
A:Accession: JC6126  
A:Molecule type: mRNA  
A:Residues: 1-347 <HAS>  
A:Cross-references: GB:066140; NID:g1628580; PIDN:AAC48667.1; PID:g1628581  
A:Note: the authors translated the codon GCC for residue 117 as Gly, GCA for residue 152  
C:Comment: This enzyme is a novel member of the histo-blood group ABO gene family. It plays a role in the differentiation of the histo-blood group I transferase  
C:Superfamily: histo-blood group I transferase  
C:Keywords: differentiation; glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 29.8%; Score 637.5; DB 2; Length 347;  
Best Local Similarity 43.5%; Pred. No. 1.1e-46;

Matches 127; Conservative 53; Mismatches 101; Indels 11; Gaps 4;

Qy 105 RIEPQLMDWPNK---NRP-DVLTVPKAPIVWEGTYDTALLLEKYATOKLTGLTVF 160  
Db 63 RSPHQ-----PKLEORPTELLTLPWADIVSEGTFNPELLOHITVPLNLTGLTVF 116

Qy 161 AVGYIEHYLEDLFLESADMYFMVGHVIRFYVIMDDTSRMPVHLNPLNPLISLOVFEIRSEKR 220  
Db 117 AVGKY-TREVFQHFLESAEQFMQGVYVYIFTNDPAGIPRVPLGPRGLLSIPIORHSR 175

Qy 221 WODISMNRKKTIGEHILAHIOHEVDFLCMDVDVQFQDNFGVETLGCOLVAQLQAWYKAS 280  
Db 176 WEEISTRMETISRIHRAHREVDYLCVDVDMVFRNPMGPEITGLDVAALHPGYAYVP 235

Qy 281 PEKTYERRELSAAYIPFEGGDFYHAAIFGCTPHTLNLTRCFCGILQDKKHIDEAQW 340  
Db 236 RQOFPYERHRTSTAFVAENEGDFYGGAVGGRVAKVYFTTCGCHMAILADKANGIMAAM 295

Qy 341 HDHSHLNKYLFLNPKTKLSPEYCHDYQIGLPSDIKSVKVAWQTKKEYNLVRN 392  
Db 296 OESHNLNRRFISHKPSKVLSPYLWDDRKPQPSLKLIRFSLDKATSWLRS 347

RESULT 9  
PC1168  
histo-blood group transferase - gorilla (fragment)  
C:Species: Gorilla gorilla (gorilla)  
C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 07-May-1999  
C:Accession: PC1168  
R:Kominato, Y.; McNeill, P.D.; Yamamoto, M.; Russell, M.; Hakomori, S.; Yamamoto, F.  
Biochem. Biophys. Res. Commun. 189, 154-164, 1992  
A:Title: Animal histo-blood group ABO genes.  
A:Reference number: PC1164; MUID:93080551  
A:Accession: PC1168  
A:Molecule type: DNA  
A:Residues: 1-189 <KOM>  
C:Superfamily: histo-blood group I transferase

Query Match 22.4%; Score 477.5; DB 2; Length 189;  
Best Local Similarity 50.3%; Pred. No. 2.1e-33;  
Matches 95; Conservative 25; Mismatches 68; Indels 1; Gaps 1;

Qy 186 RVIFYVMDDTSRMPVHLNPLHSLQVFEIRSEKRWQDISMRMKTIGEHILAHIOHEVD 245  
Db 1 RVHYVFTDQPAAVPRVTLGTGRLSVLEVYAKRWQDSMRMEMISDFCERRFLESDV 60

Qy 246 FLCMDVDVQFQDNFGVETLGCOLVAQLQAWYKASPEKTYERRELSAAYIPFEGGDFY 305  
Db 61 YLVCVDVDMFDRHVGVEILTFLCTLHPFGYSSREAFYERRPQSOAVIPKDEGDFY 120

Qy 306 HAAIFGCTPHTLNLTRCFCGILQDKKHIDEAQWHDHSHLNKYLFLNPKTKLSPEYCW 365  
Db 121 MGAFFGSGVQEVQRLTRACHQAMVDOANGIEAVWHDHSHLNKYLRLHKPTKVLSPYLW 180

Qy 366 DYQ-IGLPS 373  
Db 181 DQQLGWP 189

RESULT 10  
PC1166  
histo-blood group I transferase - chimpanzee (fragment)  
N:Alternate names: histo-blood group 2 transferase  
C:Species: Pan troglodytes (chimpanzee)  
C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 07-May-1999  
C:Accession: PC1166; PC1167  
R:Kominato, Y.; McNeill, P.D.; Yamamoto, M.; Russell, M.; Hakomori, S.; Yamamoto, F.  
Biochem. Biophys. Res. Commun. 189, 154-164, 1992  
A:Title: Animal histo-blood group ABO genes.  
A:Reference number: PC1164; MUID:93080551  
A:Accession: PC1166  
A:Molecule type: DNA

```
C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 07-May-1993
```

C:Accession: PC1169  
R:Kominato, Y.; McNeill, P.D.; Yamamoto, M.; Russell, M.; Hakomori, S.; Yamamoto, F.  
Biochem. Biophys. Res. Commun. 189, 154-164, 1992  
A:Title: Animal histo-blood group ABO genes.  
A:Reference number: PC1164; MUID:93080551  
A:Accession: PC1169  
A:Molecule type: DNA  
A:Residues: 1-189 <ROM>  
C:Superfamily: histo-blood group 1 transferase

Query Match 21.7%; Score 462.5; DB 2; Length 189;  
Best Local Similarity 48.7%; Pred. No. 3.9e-32;  
Matches 92; Conservative 26; Mismatches 70; Indels 1; Gaps 1;  
QY 186 RVIFYVMDTSTRMPVYVHLNPLHSLQVFEIRSEKRWODISMRRMKTIGEHILAHIQHEVD 245  
DB 1 RVHYVFTDPAAPVRYTLTGROLSVLGVYKRAYRWQDVSMRRMEMISDFCERRFLSEVD 60  
QY 246 FLFCMDVDQVFDQNFQVETLGQLVAQLQAWMYKASPEKFTYERRELAAAYIPFGEGDFYY 305  
DB 61 YLVCVDVDMERDVGVEILTPLFGTLHPGFGYSTREAFYERRRPSQOAYIPKDEGDFYY 120  
QY 306 HAAIFGCTPHILNLTRECFKGILODKKHIEAOHWDESHLNKYFLNKPTKILSPEYCW 365  
DB 121 LGGFFGGSVQEVQRLTRCHQAMVDAQANGIEAVVHDESHLNKYLRLRHKPTKVLSPYLV 180  
QY 366 DYQ-IGLPS 373  
DB 181 DQQLGWPA 189

RESULT 15  
PC1170  
histo-blood group 2 transferase - orangutan (fragment)  
C:Species: Pongo pygmaeus (orangutan)  
C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 07-May-1999  
C:Accession: PC1170  
R:Kominato, Y.; McNeill, P.D.; Yamamoto, M.; Russell, M.; Hakomori, S.; Yamamoto, F.  
Biochem. Biophys. Res. Commun. 189, 154-164, 1992  
A:Title: Animal histo-blood group ABO genes.  
A:Reference number: PC1164; MUID:93080551  
A:Accession: PC1170  
A:Molecule type: DNA  
A:Residues: 1-189 <ROM>  
C:Superfamily: histo-blood group 1 transferase

Query Match 21.7%; Score 462.5; DB 2; Length 189;  
Best Local Similarity 48.7%; Pred. No. 3.9e-32;  
Matches 92; Conservative 26; Mismatches 70; Indels 1; Gaps 1;  
QY 186 RVIFYVMDTSTRMPVYVHLNPLHSLQVFEIRSEKRWODISMRRMKTIGEHILAHIQHEVD 245  
DB 1 RVHYVFTDPAAPVRYTLTGROLSVLGVYKRAYRWQDVSMRRMEMISDFCERRFLSEVD 60  
QY 246 FLFCMDVDQVFDQNFQVETLGQLVAQLQAWMYKASPEKFTYERRELAAAYIPFGEGDFYY 305  
DB 61 YLVCVDVDMERDVGVEILTPLFGTLHPGFGYSTREAFYERRRPSQOAYIPKDEGDFYY 120  
QY 306 HAAIFGCTPHILNLTRECFKGILODKKHIEAOHWDESHLNKYFLNKPTKILSPEYCW 365  
DB 121 LGGFFGGSVQEVQRLTRCHQAMVDAQANGIEAVVHDESHLNKYLRLRHKPTKVLSPYLV 180  
QY 366 DYQ-IGLPS 373  
DB 181 DQQLGWPA 189

Search completed: May 10, 2002, 11:07:29  
Job time: 142 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:04:47 ; Search time 12.62 Seconds  
(without alignments)  
702.559 Million cell updates/sec

Title: US-09-863-475A-4  
Perfect score: 2136  
Sequence: 1 MITMLQDLHVNKISMSRSKS.....IKSVKVAWQTKYNLVRNV 394

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2136	100.0	394	1	US-07-914-281-4
2	2136	100.0	394	1	US-08-393-248-4
3	2136	100.0	394	1	US-08-525-058A-4
4	2136	100.0	394	2	US-08-696-731-4
5	2136	100.0	394	4	US-09-042-531-4
6	2136	100.0	394	5	PCT-US91-00899-9
7	1950	91.3	371	2	US-08-378-617A-12
8	1539.5	72.1	376	2	US-08-704-548-2
9	1515	70.9	279	5	PCT-US91-00899-8
10	1510	70.7	371	2	US-08-378-617A-10
11	1476.5	69.1	368	2	US-08-378-617A-11
12	690.5	32.3	354	1	US-07-752-101A-51
13	689	32.3	375	1	US-07-752-101A-52
14	688	32.2	353	1	US-07-752-101A-35
15	674.5	31.6	354	1	US-07-752-101A-38
16	672	31.5	353	1	US-07-752-101A-36
17	671.5	31.4	354	1	US-07-752-101A-39
18	671.5	31.4	354	1	US-07-752-101A-41
19	487	22.8	195	1	US-07-752-101A-68
20	256	12.0	154	1	US-07-752-101A-69
21	97	4.5	40	1	US-07-752-101A-10
22	97	4.5	56	1	US-07-752-101A-16
23	87.5	4.1	1720	2	US-08-477-451-12
24	82.5	3.9	822	3	US-08-941-445A-17
25	82	3.8	469	4	US-08-448-489-12
26	81	3.8	569	4	US-08-961-083-154
27	81	3.8	591	4	US-08-961-083-74

28	80.5	3.8	255	1	US-08-152-019A-36	Sequence 36, Appl
29	80.5	3.8	2213	1	US-08-727-034-3	Sequence 3, Appl1
30	80	3.7	469	3	US-08-704-711A-16	Sequence 16, Appl
31	79	3.7	758	1	US-07-756-250-16	Sequence 16, Appl
32	78.5	3.7	42	1	US-07-752-101A-8	Sequence 8, Appl1
33	77.5	3.6	2214	1	US-08-727-034-7	Sequence 7, Appl1
34	76.5	3.6	597	1	US-08-374-155A-12	Sequence 12, Appl
35	76.5	3.6	597	2	US-08-785-396-12	Sequence 12, Appl
36	75	3.5	327	3	US-09-154-874-8	Sequence 8, Appl1
37	75	3.5	462	4	US-08-068-392-3	Sequence 3, Appl1
38	75	3.5	462	4	US-08-396-988-3	Sequence 3, Appl1
39	75	3.5	485	3	US-08-600-658-7	Sequence 7, Appl1
40	75	3.5	485	4	US-09-170-670-6	Sequence 6, Appl1
41	75	3.5	485	4	US-09-193-068-6	Sequence 6, Appl1
42	75	3.5	485	4	US-09-183-412-6	Sequence 6, Appl1
43	75	3.5	485	4	US-09-354-191A-7	Sequence 7, Appl1
44	75	3.5	2332	1	US-07-864-004B-4	Sequence 4, Appl1
45	75	3.5	2332	1	US-08-251-937A-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1  
US-07-914-281-4  
; Sequence 4, Application US/07914281  
; Patent No. 5324663  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/914,281  
; FILING DATE: 19920720  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 394 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-07-914-281-4

Query Match 100.0% Score 2136; DB 1; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.4e-214;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSEKIMNVKGKVLMLLIVSTVVVVFWEY 60

```
Db 1 MITLQDLHVNKISMSRSKSETSLPSSSGSGOEKIMNVKGVILLMLLIVSTVVVWFWEYV 60
Qy 61 NRIPEVGENRQKQDWPFQSWFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLWDFNPKNR 120
Db 61 NRIPEVGENRQKQDWPFQSWFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLWDFNPKNR 120
Qy 121 PDVLTVPWKAPIVWEGTYDTALLEKYVATQKLTGLTVFAVGKVIHYHLEDFLESADMY 180
Db 121 PDVLTVPWKAPIVWEGTYDTALLEKYVATQKLTGLTVFAVGKVIHYHLEDFLESADMY 180
Qy 181 FMVGHRIYFYVMIDTSTRMPVYVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240
Db 181 FMVGHRIYFYVMIDTSTRMPVYVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240
Qy 241 QHEYDFLFCMDVDQVFDNFGVETLGLVAQLQAWWYKASPEKTYERRELSAAIYPFGE 300
Db 241 QHEYDFLFCMDVDQVFDNFGVETLGLVAQLQAWWYKASPEKTYERRELSAAIYPFGE 300
Qy 301 GDFYHAAIFGGTPTTHILNLTRECFKGLQDKKHDIEAQWHDHSHLNKYFLFNKPTKILS 360
Db 301 GDFYHAAIFGGTPTTHILNLTRECFKGLQDKKHDIEAQWHDHSHLNKYFLFNKPTKILS 360
Qy 361 PEYCDYQIGLPSDIKSVKVAWQKEYNLVRNV 394
Db 361 PEYCDYQIGLPSDIKSVKVAWQKEYNLVRNV 394

RESULT 2
US-08-393-246-4
; Sequence 4, Application US/08393246
; Patent No. 5595900
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,246
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
```

```
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-393-246-4

Query Match 100.0%; Score 2136; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.4e-214;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MITLQDLHVNKISMSRSKSETSLPSSSGSGOEKIMNVKGVILLMLLIVSTVVVWFWEYV 60
Db 1 MITLQDLHVNKISMSRSKSETSLPSSSGSGOEKIMNVKGVILLMLLIVSTVVVWFWEYV 60
Qy 61 NRIPEVGENRQKQDWPFQSWFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLWDFNPKNR 120
Db 61 NRIPEVGENRQKQDWPFQSWFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLWDFNPKNR 120
Qy 121 PDVLTVPWKAPIVWEGTYDTALLEKYVATQKLTGLTVFAVGKVIHYHLEDFLESADMY 180
Db 121 PDVLTVPWKAPIVWEGTYDTALLEKYVATQKLTGLTVFAVGKVIHYHLEDFLESADMY 180
Qy 181 FMVGHRIYFYVMIDTSTRMPVYVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240
Db 181 FMVGHRIYFYVMIDTSTRMPVYVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240
Qy 241 QHEYDFLFCMDVDQVFDNFGVETLGLVAQLQAWWYKASPEKTYERRELSAAIYPFGE 300
Db 241 QHEYDFLFCMDVDQVFDNFGVETLGLVAQLQAWWYKASPEKTYERRELSAAIYPFGE 300
Qy 301 GDFYHAAIFGGTPTTHILNLTRECFKGLQDKKHDIEAQWHDHSHLNKYFLFNKPTKILS 360
Db 301 GDFYHAAIFGGTPTTHILNLTRECFKGLQDKKHDIEAQWHDHSHLNKYFLFNKPTKILS 360
Qy 361 PEYCDYQIGLPSDIKSVKVAWQKEYNLVRNV 394
Db 361 PEYCDYQIGLPSDIKSVKVAWQKEYNLVRNV 394

RESULT 3
US-08-525-058A-4
; Sequence 4, Application US/08525058A
; Patent No. 5770420
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,058A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
```



```

; ; CURRENT APPLICATION DATA:
; ; APPLICATION NUMBER: US/08/696,731
; ; FILING DATE: 14-AUG-1996
; ; CLASSIFICATION: 435
; ; PRIOR APPLICATION DATA:
; ; APPLICATION NUMBER: 08/393,246
; ; FILING DATE:

```

PRODUCTS FOR THE SYNTHESIS  
CHARIDE STRUCTURES ON GLYCOPROTEINS,  
, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES

LELLAND, MAIER & NEUSTADT,  
Highway, Fourth Floor

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/042,531  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,246  
FILING DATE: US 08/220,433  
FILING DATE: 30-MAR-1994  
APPLICATION NUMBER: US 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-042-531-4

Query Match 100.0%; Score 2136; DB 4; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.4e-214;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVWFWEYV 60  
DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVWFWEYV 60

QY 61 NRIPEVGENRWQDMWPFPSWFKNGTHSYQEDNVGREGKRGNGDRIEPEQLWDMFNPKNR 120  
DB 61 NRIPEVGENRWQDMWPFPSWFKNGTHSYQEDNVGREGKRGNGDRIEPEQLWDMFNPKNR 120

QY 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKVIHYLEDLFLESADMY 180  
DB 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKVIHYLEDLFLESADMY 180

QY 181 FMVGHVRVIFYVMIDTSTRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
DB 181 FMVGHVRVIFYVMIDTSTRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240

QY 241 QHEVDLFCMDVDQVDFQDNFVETLGQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGE 300  
DB 241 QHEVDLFCMDVDQVDFQDNFVETLGQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGE 300

QY 301 GDFYVHAAIFGGTPTTHLNLTRCFKGILODKKHIDIEAQWHDHSHLNKYFLFNKPTKILS 360  
DB 301 GDFYVHAAIFGGTPTTHLNLTRCFKGILODKKHIDIEAQWHDHSHLNKYFLFNKPTKILS 360

QY 361 PEYCDWYQIGLPSDIKSVKVAWQTKYENLVNRNV 394  
DB 361 PEYCDWYQIGLPSDIKSVKVAWQTKYENLVNRNV 394

RESULT 6  
PCT-US91-00899-9  
Sequence 9, Application PC/TUS9100899  
GENERAL INFORMATION:  
APPLICANT: Lowe, John B.  
TITLE OF INVENTION: Method and Products For the Synthesis of  
TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,

TITLE OF INVENTION: Or as Free Molecules, and For the Isolation of Cloned  
TITLE OF INVENTION: Genetic Sequences That Determine These Structures  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/00899  
FILING DATE: 19910214  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye Ph.D., Jean-Paul  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-021-55 PCT  
TELEPHONE: (703)521-5940  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
TISSUE TYPE: Blood  
PCT-US91-00899-9

Query Match 100.0%; Score 2136; DB 5; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.4e-214;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVWFWEYV 60  
DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVWFWEYV 60

QY 61 NRIPEVGENRWQDMWPFPSWFKNGTHSYQEDNVGREGKRGNGDRIEPEQLWDMFNPKNR 120  
DB 61 NRIPEVGENRWQDMWPFPSWFKNGTHSYQEDNVGREGKRGNGDRIEPEQLWDMFNPKNR 120

QY 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKVIHYLEDLFLESADMY 180  
DB 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKVIHYLEDLFLESADMY 180

QY 181 FMVGHVRVIFYVMIDTSTRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
DB 181 FMVGHVRVIFYVMIDTSTRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240

QY 241 QHEVDLFCMDVDQVDFQDNFVETLGQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGE 300  
DB 241 QHEVDLFCMDVDQVDFQDNFVETLGQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGE 300

QY 301 GDFYVHAAIFGGTPTTHLNLTRCFKGILODKKHIDIEAQWHDHSHLNKYFLFNKPTKILS 360  
DB 301 GDFYVHAAIFGGTPTTHLNLTRCFKGILODKKHIDIEAQWHDHSHLNKYFLFNKPTKILS 360

QY 361 PEYCDWYQIGLPSDIKSVKVAWQTKYENLVNRNV 394  
DB 361 PEYCDWYQIGLPSDIKSVKVAWQTKYENLVNRNV 394

```

RESULT 7
US-08-378-617A-12
: Sequence 12, Application US/08378617A
: Patent No. 5849991
: GENERAL INFORMATION:
: APPLICANT: d'Apice, Anthony J.F.
: APPLICANT: Pearse, Martin J.
: APPLICANT: Robins, Allan J.
: APPLICANT: Crawford, Robert J.
: APPLICANT: Rathjen, Peter D.
: TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF
: TITL OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 120 South Sixth Street, Suite 2500
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30B
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/378.617A
: FILING DATE: 26-JAN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Ellinger, Mark S.
: REGISTRATION NUMBER: 34,812
: REFERENCE/DOCKET NUMBER: 06868/005001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (612) 335-5070
: TELEFAX: (612) 288-9696
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 371 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-378-617A-12

```

	Query Match	91.3%	Score 1950;	DB 2;	Length 371;
	Best Local Similarity	96.5%	Pred. No. 3.2e-195;		
	Matches 358; Conservative	1;	Mismatches 0;	Indels 12;	Gaps 1;
Qy	36	MNVGKGVTLMLIVSTVVVFWFXYVN-----RIPEVGENRRKQKDWPFWSFKN	83		
				:	
Dd	1	MNVGKGVTLMLIVSTVVVFWFXYVNSPDGSFLWJYHTKIPEVGENRRKQKDWPFWSFKN	60		
Qy	84	GTHSYQEDNVGREGKRGNGRDRIEPPQLWDFNFKNRPDVLTVTPWKAPIVWEGTYDTAL	143		
Dd	61	GTHSYQEDNVGREGKRGNGRDRIEPPQLWDFNFKNRPDVLTVTPWKAPIVWEGTYDTAL	120		
Qy	144	LEKYATQKLTVGLTVFAVGKYYIEHYLEDFFLESADMYFMVGHRTVFYVMIDTTSRMPVVH	203		
Dd	121	LEKYATQKLTVGLTVFAVGKYYIEHYLEDFFLESADMYFMVGHRTVFYVMIDTTSRMPVVH	180		
Qy	204	LNP LHSLOVFEIRSEKRWDIISMRRMKTIGEHILAHIOHEVDFFLCMDVDQVFODNFGVE	263		
Dd	181	LNP LHSLOVFEIRSEKRWDIISMRRMKTIGEHILAHIOHEVDFFLCMDVDQVFODNFGVE	240		
Qy	264	TLGOLVAOLOAWMYKASPEKFTYERRELSAAYIPFGEGRDYFYHAAIFGGTPTHILNLTRE	323		
Dd	241	TLGOLVAOLOAWMYKASPEKFTYERRELSAAYIPFGEGRDYFYHAAIFGGTPTHILNLTRE	300		
Qy	324	CFKGILQDKKHDIIEAQWHDESHLNKPYFLFNPKPTKILSP EYCDWYOIGLPSDIKS VKVAMQ	383		
Dd	301	CFKGILQDKKHDIIEAQWHDESHLNKPYFLFNPKPTKILSP EYCDWYOIGLPSDIKS VKVAMQ	360		

```

Qy 384 TKEYNLVRNNV 394
      |||||
Db 361 TKEYNLVRNNV 371

RESULT 8
US-08/704-548-2
: Sequence 2, Application US/08704548
: Patent No. 5879675
: GENERAL INFORMATION:
: APPLICANT: GALILI, URI
: APPLICANT: REPIK, PATRICIA M.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR VACCINES
: TITLE OF INVENTION: COMPRISING ALPHA-GALACTOSYL EPTIOPES
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
: STREET: Suite 1800, Two Penn Center Plaza
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19102
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/704,548
: FILING DATE: 11-SEP-1996
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Monaco, Daniel A.
: REGISTRATION NUMBER: 30,480
: REFERENCE/DOCKET NUMBER: 8760-2 CII
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (215) 568-8383
: TELEFAX: (215) 568-5549
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 376 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08/704-548-2

Query Match 72.1%; Score 1539.5; DB 2; Length 376;
Best Local Similarity 72.6%; Pred. No. 2.2e-152;
Matches 273; Conservative 43; Mismatches 43; Indels 17; Gaps

Qy 36 MNVKGKVTLLMLVSTVVVVFWEVNN-----RIPEGVENRWOKDMWFPSPKFN 73
      |||||
Db 1 MNVKGKVTLLMLVSTVVVVFWEVNNSPESFLWIYHSKNPDESSAOKDMWFPSPKFN 70

Qy 84 GTHSYQE-----DNVEGRREKGRNGDIEBPQLMDWPNKRPDVLTVTPWKAPIVWEGT 138
      |||||
Db 61 GIHNYQEEEDTDKEKGREEEQKEDDTTELRLMDWPNPKKREVTVTQWKAPVWEGT 120

Qy 139 YDTALLKKYATQKLTVGLTVFAVGKVIHYLEDFLSADMYFMVGHVRVIFYVMIDDTSR 198
      |||||
Db 121 YNKAILLENYAKQKITVGLTVFAIGRVIEHYLEEFVTSANRYFMVGHKVIIFYVMVDVSK 180

Qy 199 MPVVHLNPLHSLOVFEJLRSERKRWODISMRRMKTIGEHILAHIOHEVDFLFCMDVDQVFO 258
      |||||
Db 181 APFTLEGLPLRSFKFYKPKRWODISMRRMKTIGEHILAHIOHEVDFLFCMDVDQVFO 240

Qy 259 NFGVETLQGLVAQLQAWMYKASPEKFTYERRELSAAYIPEGGEFYHHAIFGCTPTHIL 318
      |||||
Db 241 HFGVETLQGSVAQLQAWMYKADPDFTYERRKSSAAYIPEGGDFYHHAIFGCTPIQVL 300

Qy 319 NLTRECFKGIQLDQKKHDI EAQWHDESHNLKYFLNPKRTILSPYECWDYTOIGLPSDKSV 378
      |||||

```

Query Match	72.1%	Score 1539.5	DB 2	Length 376
Best Local Similarity	72.6%	Pred. No. 2.2e-152		
Matches	273	Conservative .43	Mismatches 43	Indels 17
Gaps	2			

  

y	36	MYNGKGVILLMLIVSTVVVVFEYVN-----RIPEGENRWQKQWFFSWPKN	83
b	1	:    :        :    ::	60
y	84	GTHSYQE----DNVSGRREKGRGNDRIEPPQLWQWFNPKNRPDVLTVTPWKAPLWVST	138
b	61	:    :          :          :          :	120
y	139	YDFTALLEKYYATOKLTVGLTVFAVGKY IEHYLEDPLFESADMYFMVGHGRVIFYVVMIDDTSR	198
b	121	:    :          :          :          :          :	180
y	199	MPVHLNPLHSLOVFEIRSEKRWQDLSMMRMTIGSHILAHQHEVDLFCLMDVQVQFD	258
b	181	:    :          :          :          :          :	240
y	259	NFGVETLGQALQALQAWMYKASPEKFTYERRELSAAYIPFGEGDFYVYHAAIEGGTPHIL	318
b	241	:    :          :          :          :          :	300
y	319	NLTRECFCGLQPKHHDIEAQWHDESHNLNKFYLFNKPTKILSPCYCWDYQIGLPSDIKS	378
b		:    :          :          :          :          :	



299 QECFKGILQDKNDIEAEMHDESHLNKYLFLNKPTKILSPYCWYHIGMSVDIRVIA 358  
382 WOTKEYNLVRNV 394  
359 WQKEYNLVRNNI 371

RESULT 11  
US-08-378-617A-11  
: Sequence 11, Application US/08378617A  
: Patent No. 5849991  
: GENERAL INFORMATION:  
: APPLICANT: d'Apice, Anthony J.F.  
: APPLICANT: Pearce, Martin J.  
: APPLICANT: Robins, Allan J.  
: APPLICANT: Crawford, Robert J.  
: APPLICANT: Rathjen, Peter D.  
: TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF  
: TITLE OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION  
: NUMBER OF SEQUENCES: 33  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Fish & Richardson  
: STREET: 120 South Sixth Street, Suite 2500  
: CITY: Minneapolis  
: STATE: MN  
: COUNTRY: USA  
: ZIP: 55402  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent In Release #1.0, Version #1.30B  
: CURRENT APPLICATION NUMBER: US/08/378,617A  
: FILING DATE: 26-JAN-1995  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Ellinger, Mark S.  
: REGISTRATION NUMBER: 34,812  
: REFERENCE/DOCKET NUMBER: 06868/005001  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (612) 335-5070  
: TELEFAX: (612) 288-9696  
: INFORMATION FOR SEQ ID NO: 11:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 368 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-378-617A-11

Query Match 69.1%; Score 1476.5; DB 2; Length 368;  
Best Local Similarity 72.2%; Pred. No. 8e-146;  
Matches 268; Conservative 41; Mismatches 47; Indels 15; Gaps 3;

QY 36 MNVKGKVLMLVSTVIVVFWFVYV-----NRIPYGENRWKQDMWFPWFKN 83  
DB 1 MNVKGKVLMLVSTVIVVFWFVYVYHSPECSLEFWNPSRNPVEVGGSIQKGMWLPWFNN 60

QY 84 GTHSYQEDNVNVEGRREKRGDRIEEPQLMDWPNKRPDVLVTPWKPAPIWEGTYDTAL 143  
DB 61 GYH--BEDGDINEEKQORNEDE-SKLKLSDFWENPFKRPVEVVTWKWKPAPVWEGTYNRAV 117

QY 144 LEKYYATOKLTGLTVFAVGKYTEHYLEDLESADMYFMVGHVRYFYVMDTTSRMPVVH 203  
DB 118 LDNYAKQKTLTVGLTVFAVGRYTEHYLEEFLLTSANKHFMVGHVRYFYIMVDDYSRMLPIE 177

QY 204 LNPILHSQVPEIRSEKRWODISMWRKKTIGEHLAHQIEHVDVDFLCMDVDVQDFQDNFGVE 263  
DB 178 LGPURSEKVPKIPKRWQODISMWRKKTIGEHLAHQIEHVDVDFLCMDVDVQDFQDNFGVE 237

QY 264 TIGOLVAQLQAMWYKASPEKFTYERRELSAAYIPFGEQGFYHAAIFGGTPTTHILNLTRE 323  
DB 238 TLGESVAQLQAMWYKADPNDFTYERRKESAAYPFGEQGFYHAAIFGGTPTQVLNITQE 297

QY 324 CFKGILODKKHDEIAQWHDHSHLNKYLFLNKPTKILSPYCWYQIGLPSDIKSVKVAWQ 383  
DB 298 CFKGILODKKHDEIAQWHDHSHLNKYLFLNKPTKILSPYCWYHYHGLPADIKLVKMSWQ 357

QY 384 TREYNLVRNV 394  
DB 358 TREYNVVRNV 368

RESULT 12  
US-07-752-101A-51  
: Sequence 51, Application US/07752101A  
: Patent No. 5326857  
: GENERAL INFORMATION:  
: APPLICANT: Yamamoto, Fum1-ichiro  
: APPLICANT: White, Thayer  
: APPLICANT: Hakomori, Sen-itiroh  
: APPLICANT: Clausen, Henrik  
: TITLE OF INVENTION: ABO GENOTYPING  
: NUMBER OF SEQUENCES: 69  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Seed and Berry  
: STREET: 6300 Columbia Center, 701 Fifth Avenue  
: CITY: Seattle  
: STATE: Washington  
: COUNTRY: U.S.  
: ZIP: 98104  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent In Release #1.0, Version #1.25  
: CURRENT APPLICATION NUMBER: US/07/752,101A  
: FILING DATE: 19910829  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Sharkey, Richard G.  
: REGISTRATION NUMBER: 32,629  
: REFERENCE/DOCKET NUMBER: 150036.406C1  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 206-622-4900  
: TELEFAX: 206-682-6031  
: TELEX: 3723836  
: INFORMATION FOR SEQ ID NO: 51:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 354 amino acids  
: TYPE: AMINO ACID  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: peptide  
: HYPOTHETICAL: NO  
: ANTI-SENSE: NO  
: FRAGMENT TYPE: N-terminal  
US-07-752-101A-51

Query Match 32.3%; Score 690.5; DB 1; Length 354;  
Best Local Similarity 45.8%; Pred. No. 7.5e-64;  
Matches 142; Conservative 47; Mismatches 108; Indels 13; Gaps 4;

QY 95 GRREKG-----RNGDRIEEPQLMDWPNKPN-----RPDVLVTPWKPAPIWEGTYDTAL 143  
DB 45 GSLERGFCAVAVRDPDLQVSLPRMYPQPKVLTPCKRQDLVLTWLPAPIWEGTFNIDI 104

QY 144 LEKYYATOKLTGLTVFAVGKYTEHYLEDLESADMYFMVGHVRYFYVMDTTSRMPVVH 203  
DB 105 LNFQRLQNTTIGLTVFAIKKYA-FUKLFLTAERHFVGHVRYFYVFTDQPAVAPRV 163



Qy	95	GREKG-----RNGDRIEPPQLWDWNPKN-----RPDVLTVTPMKAPIVWEGTYDTAL	143
Db	45	GSLERCFCMAVREPDHLQRLVSFLPRMVPQPKVLTPCRKDVLPVVTPLAPIVWEGTFNIDI	104
Qy	144	LEKYATQKLTVGLTVFVNGKYTEHLYEDEFESADMYFVWGIRVIYFYVIMDITSRMPVVH	203
Db	105	LNEQFLQNTTIGLTVFAIKKYA-FLKFLETAERKFMVGRHYHYVFTDQPAAPRVY	163
Qy	204	LNP LHSLOVEIRSEKRWODISMRMKTTCHEHTLAHQHEVDPLFCMDVDVDFODNFGVE	263
Db	164	LGTGRQLSVLEVRAYRKWDVSMRRMEMISDFCERREFLSEVDYLVLCVDVMEFRDHVGVE	223
Qy	264	TLQQLVAQLQAMWYKASPEKTFVERRELSAAYIPFCEGDFYFHAALFGGTPTTHLNLTR	323
Db	224	ILTPLECTLHPGFIYGSREAFTEYRPSQAYLPKDEGDFYLLGGFFGSGVQVEQLRTA	283
Qy	324	CFKGILQDKKHDIQAQWHIDESHLNKFYLFNPKPTKILSPCYCWDQ-IGLPSDIKSVKAW	382
Db	284	CHQAMVVDQANGIEAVWHIDESHLNLYLRHKPTKVLSPVLYMDQQLGMPAVLRLKLFRTA	343
Qy	383	QTKEYNLVRN	392
Db	344	VPKNHOAVRN	353

```

1 GENERAL INFORMATION:
2 APPLICANT: Yamamoto, Fumi-Ichiro
3 APPLICANT: White, Thayer
4 APPLICANT: Hakomori, Sen-itiroh
5 APPLICANT: Clausen, Henrik
6 TITLE OF INVENTION: ABO GENOTYPING
7 NUMBER OF SEQUENCES: 69
8 CORRESPONDENCE ADDRESS:
9 ADDRESSEE: Seed and Berry
10 STREET: 6300 Columbia Center, 701 Fifth Avenue
11 CITY: Seattle
12 STATE: Washington
13 COUNTRY: U.S.
14 ZIP: 98104
15
16 COMPUTER READABLE FORM:
17 MEDIUM TYPE: Floppy disk
18 COMPUTER: IBM PC Compatible
19 OPERATING SYSTEM: PC-DOS/MS-DOS
20 SOFTWARE: PatentIn Release 1.0, Version 1.2.25
21
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/07/752,101A
24 FILING DATE: 19910829
25 CLASSIFICATION: 435
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Sharkey, Richard G.
28 REGISTRATION NUMBER: 32,629
29 REFERENCE/DOCKET NUMBER: 150036.406C1
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 206-622-4900
32 TELEFAX: 206-682-6031
33 TELEX: 3723836
34
35 INFORMATION FOR SEQ ID NO: 38:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 354 amino acids
38 TYPE: AMINO ACID
39 STRANDEDNESS: single
40 TOPOLOGY: linear
41 MOLECULE TYPE: peptide
42 HYPOTHETICAL: NO
43 ANTI-SENSE: NO
44 FRAGMENT TYPE: N-terminal
45
46 US-07-752-101A-38

```

Query Match 31.6%; Score 674.5; DB 1; Length 354;  
Best Local Similarity 45.2%; Pred. NO. 3.5e-62;  
Matches 140; Conservative 47; Mismatches 110; Indels 13; Gaps 4;





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:01:17 ; Search time 25.31 Seconds  
(without alignments)  
1153.097 Million cell updates/sec

Title: US-09-863-475A-4

Perfect score: 2136

Sequence: 1 MITMLQDLHVNKISMSRSK.....IKSVKVAQTKYENLVNRY 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2136	100.0	394	12 AAR13750	GDP-Fuc:(beta-D-Ga
2	2136	100.0	394	15 AAR45935	A glycosyltransfer
3	2136	100.0	394	18 AAW13639	Murine alpha(1,3)-
4	1539.5	72.1	376	16 AAR80016	Marmoset alpha-1,3
5	1532	71.7	363	19 AAW49687	Porcine alpha-1,3-
6	1525	71.4	359	15 AAR62508	Galactosyl transfe
7	1525	71.4	359	17 AAR90573	pig alpha(1,3)-gal
8	1513	70.8	375	19 AAW49686	Porcine alpha-1,3-
9	1500	70.2	371	16 AAR85082	Porcine alpha (1,3
10	1493	69.9	371	16 AAR76777	Pig alpha-1,3-gala
11	1431.5	67.0	354	19 AAR49688	Porcine alpha-1,3-

12	1428.5	66.9	342	19 AAW49689	Porcine alpha-1,3-
13	1215	56.9	313	15 AAR62507	Galactosyl transfe
14	690.5	32.3	335	22 AAE05193	Human drug metabol
15	690.5	32.3	354	12 AAR11789	Histo-blood gp. A
16	690.5	32.3	354	15 AAR57020	Human A transferas
17	689	32.3	375	15 AAR57021	Human A transferas
18	688	32.2	353	12 AAR11317	Histo-blood gp. A
19	688	32.2	353	12 AAR57011	Histo-blood gp. A
20	687.5	32.2	354	12 AAR11790	Human A transferas
21	687.5	32.2	354	12 AAR11792	Histo-blood gp. B
22	674.5	31.6	354	15 AAR57013	Human A transferas
23	673.5	31.5	358	15 AAR57014	Human B transferas
24	673	31.5	347	22 AAM39251	Human polypeptide
25	671.5	31.4	354	15 AAR57016	Human B transferas
26	646	30.2	402	22 AAM41037	Human polypeptide
27	487	22.8	195	15 AAR57024	Partial sequence o
28	256	12.0	154	15 AAR57025	hgt4 showing homol
29	230	10.8	149	22 AAM25860	Human protein sequ
30	229	10.7	106	21 AAY86491	Human gene 59-enco
31	216	10.1	100	20 AAY26039	Secreted protein n
32	213	10.0	100	22 AAB73508	Human transferase
33	175	8.2	90	21 AAY86273	Human secreted pro
34	175	8.2	90	21 AAY86323	Human secreted pro
35	175	8.2	90	21 AAY86489	Human gene 59-enco
36	97	4.5	40	15 AAR57004	N-terminal fragmen
37	97	4.5	56	15 AAR57010	N-terminal fragmen
38	97	4.5	432	22 AAG91438	C glutamicum prote
39	97	4.5	432	22 AAB79092	Corynebacterium gl
40	97	4.5	432	22 AAB79725	Corynebacterium gl
41	93.5	4.4	334	22 AAG81594	S. epidermidis ope
42	89.5	4.2	513	14 AAR39900	21B4/rhoptry prote
43	88.5	4.1	443	21 AAG13394	Arabidopsis thalia
44	88.5	4.1	460	21 AAG13393	Arabidopsis thalia
45	88.5	4.1	468	21 AAG13392	Arabidopsis thalia

ALIGNMENTS

RESULT 1  
AAR13750  
ID AAR13750 standard; Protein; 394 AA.  
XX  
AC AAR13750;  
DT 07-NOV-1991 (first entry)  
XX  
DE GDP-Fuc:(beta-D-Gal(1,4/1,3))-D-GlcNAc(/Glc)alpha(1,3/1,4)  
DE -fucosyltransferase.  
XX  
KW Glycosyltransferase.  
XX  
OS Mus musculus.  
XX  
PN WO9112340-A.  
XX  
PD 22-AUG-1991.  
XX  
PF 14-FEB-1991; 91WO-US00899.  
XX  
PR 12-DEC-1990; 90US-0627621.  
PR 14-FEB-1990; 90US-0479858.  
PR 14-FEB-1990; 90US-0480133.  
XX  
XX (UNMI ) UNIV OF MICHIGAN.  
PI Lowe JB;  
XX WPI; 1991-267151/36.  
DR N-PSDB; AAG13331.  
XX  
XX Isolation of gene conveying post-translational characteristic -  
PT e.g. the presence of soluble or membrane bound oligo or

PT polysaccharide or glycosyltransferase.

XX Disclosure; Fig 2; 155pp; English.

CC The amino acid sequence codes for a protein capable of functioning  
CC as a UDP- Gal: [beta-D-Gal(1,4)]-D-GlcNAc alpha (1,3)galacto-  
CC syltransferase.. The products of this enzyme, sub-terminal alpha  
CC (1,3) and alpha(1,4) fucose residues are used in the post-  
CC translational modification of the oligosaccharides on cell-surface,  
CC intracellular or secreted proteins or lipids. These can be used for  
CC the prodn. of diagnostics and therapeutics. There is a single  
CC transmembrane domain consisting of a 19 amino acid hydrophobic  
CC segment flanked by basic residues and a large (presumably  
CC catalytic) C-terminal domain that would ultimately be targeted to  
CC the lumen of the Golgi. It has two potential N-glycosylation sites  
CC indicating that as with other glycosyltransferases, it may be  
CC synthesised as a glycoprotein. It is representative of a Type II  
CC transmembrane protein. See also AAR13749-R13752.

XX Sequence 394 AA;

Query Match 100.0%; Score 2136; DB 12; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1e-206;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVTVVWFWEYV 60  
DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVTVVWFWEYV 60  
QY 61 NRPEVGENRQKDWPPSFKNKTHSYQEDNVEGRKGRNGDRIEEPQLWDFNPKNR 120  
DB 61 NRPEVGENRQKDWPPSFKNKTHSYQEDNVEGRKGRNGDRIEEPQLWDFNPKNR 120  
QY 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTVGTVFAVGKYIEHYLEDLFLESADMY 180  
DB 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTVGTVFAVGKYIEHYLEDLFLESADMY 180  
QY 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
DB 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
QY 241 QHEVDLFCHDVOVFDQNFQVETLGLVAQLQAWWYKASPEKFTYERRELSAAYIPFGE 300  
DB 241 QHEVDLFCHDVOVFDQNFQVETLGLVAQLQAWWYKASPEKFTYERRELSAAYIPFGE 300  
QY 301 GDFYHAAIFGGTPTHTLNLTRCFKGILODKKHDIQAWHDESHLNKYFLFNKPTKILS 360  
DB 301 GDFYHAAIFGGTPTHTLNLTRCFKGILODKKHDIQAWHDESHLNKYFLFNKPTKILS 360  
QY 361 PEYCDWYQIGLPSDIKSVKVAWQTKKEYNLVRNV 394  
DB 361 PEYCDWYQIGLPSDIKSVKVAWQTKKEYNLVRNV 394

RESULT 2

AAR45935  
ID AAR45935 standard; Protein; 394 AA.

XX AAR45935;

XX 26-JUL-1994 (first entry)

DT A glycosyltransferase.

XX Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;  
XX surface; oligosaccharide.

OS Homo sapiens.

XX WO9402616-A.

XX 03-FEB-1994.

PD

XX 20-JUL-1993; 93WO-US06703.

XX 20-JUL-1992; 92US-0914281.

XX (UNMI ) UNIV MICHIGAN.

XX Lowe JB;

XX WPI; 1994-048874/06.

XX N-PSDB; AAQ56907.

XX DNA fragment encoding a glycosyltransferase - can be used for in  
XX vitro reactions to modify cell surface oligosaccharides) e.g.  
XX blood gp. determinants, to protect against transplant rejection  
XX  
XX Disclosure; Fig 2; 249pp; English.  
XX  
XX The sequence is that of a human glycosyl transferase. The enzyme  
XX may be non glycosylated. This prevents premature loss of enzyme  
XX activity. It can also be used in in vitro reactions to modify cell  
XX surface oligosaccharide mols. e.g. blood group determinants.  
XX See also AAR45933-9.

XX Sequence 394 AA;

Query Match 100.0%; Score 2136; DB 15; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1e-206;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVTVVWFWEYV 60  
DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVTVVWFWEYV 60  
QY 61 NRPEVGENRQKDWPPSFKNKTHSYQEDNVEGRKGRNGDRIEEPQLWDFNPKNR 120  
DB 61 NRPEVGENRQKDWPPSFKNKTHSYQEDNVEGRKGRNGDRIEEPQLWDFNPKNR 120  
QY 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTVGTVFAVGKYIEHYLEDLFLESADMY 180  
DB 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTVGTVFAVGKYIEHYLEDLFLESADMY 180  
QY 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
DB 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
QY 241 QHEVDLFCHDVOVFDQNFQVETLGLVAQLQAWWYKASPEKFTYERRELSAAYIPFGE 300  
DB 241 QHEVDLFCHDVOVFDQNFQVETLGLVAQLQAWWYKASPEKFTYERRELSAAYIPFGE 300  
QY 301 GDFYHAAIFGGTPTHTLNLTRCFKGILODKKHDIQAWHDESHLNKYFLFNKPTKILS 360  
DB 301 GDFYHAAIFGGTPTHTLNLTRCFKGILODKKHDIQAWHDESHLNKYFLFNKPTKILS 360  
QY 361 PEYCDWYQIGLPSDIKSVKVAWQTKKEYNLVRNV 394  
DB 361 PEYCDWYQIGLPSDIKSVKVAWQTKKEYNLVRNV 394

RESULT 3

AAR13639  
ID AAR13639 standard; Protein; 394 AA.

XX AAR13639;

XX 19-JUN-1997 (first entry)

XX Murine alpha(1,3)-galactosyltransferase.

XX Alpha(1,3)-galactosyltransferase; glycosylation; oligosaccharide.

XX Mus sp.

```

XX PN WO9709421-A1.
XX PD 13-MAR-1997.
XX PF 06-SEP-1996; 96WO-US13816.
XX PR 08-SEP-1995; 95US-0525058.
XX PA (UNMI ) UNIV MICHIGAN.
XX PI Legault DJ, Lowe JB;
XX PWPI; 1997-192897/17.
XX DR N-PSDB; AAT61676.
XX PT New recombinant fucosyltransferase proteins - useful for modifying
XX PT cell surface oligosaccharide structures
XX PS Example 2; Page 272-274; 329pp; English.
XX CC Murine UDP-Gal:beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-
XX CC galactosyltransferase (AAW13639) catalyses the a transglycosylation
XX CC reaction between UDP-Gal and N-acetyllactosamine and is associated
XX CC with surface-localised expression of Gal(alpha1-3)Gal linkages.
XX CC Its amino acid sequence was deduced from a cDNA clone (AAT61676)
XX CC obt'd. by transfecting COS-1 cells with cDNA derived from mouse F9
XX CC teratocarcinoma cells, and screening the transfected cells for
XX CC surface-localised Gal(alpha1-3)Gal linkages. When expressed in
XX CC animal cell lines, the enzyme provides specific capabilities with
XX CC respect to post-translational modification of the oligosaccharides
XX CC of expressed proteins or lipids. The enzyme can also be used to
XX CC raise antibodies and to screen for inhibitor cpds.
XX CC Sequence 394 AA;
XX SQ
Query Match 100.0%; Score 2136; DB 18; Length 394;
Best Local Similarity 100.0%; Pred. No. 1e-206;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MITMLQDLHVNKLSMSRSKSETSLPSSRSGSOEIKMNVKGVILLMLIVSTVVVFWEYV 60
DB 1 mitmlqdlhvnklsmsrsksetslpsrsgsqekimnvkgvillmlivstvvvfweyv 60
QY 61 NRPEVGENRQKDWMPSPFKNKTHSYQEDNVGREGKRGNDRIEPPQLWDFWPNKNR 120
DB 61 nrpevgenrwdkdwmpsfknkthsyqednvgrregkrngndrieepqlwdfwnpknr 120
QY 121 PDVLTVPWKAPVWEGTYDTALLEKYATOKLTVGLTVFAVGKYIEHYLEDLESADMY 180
DB 121 pdvltvtpwkaplvwegtydtallekyatqkltvgtlvfavgkylehyledfilesadmy 180
QY 181 FMVGHVRVIFVYVIMDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHLIAHI 240
DB 181 fmvghrvifvymddtsrmpvvhnlplhslqvfseirsekrdwdismmrmktigehlliahi 240
QY 241 QHEVDLFLCNDVQVDFONFVETLQGLVAQLQAWMYKASPEKFTYERRELSAAIIPFGE 300
DB 241 qhevdfllcndvqvfdfonfvettlqglvqlqawmykaspekftyerrrelsaaiipfge 300
QY 301 GDFYHAAIFGGPTTHILNLTRECFCGILQDKKHDIQAQWHDSEHLNKNYFLFNKPTKILS 360
DB 301 gdfyyhaaifggptthilnltrecfcgilkdkkhdieaqwhdeshlnknflynkptkils 360
QY 361 PEYCDWQIGLPSDQIKSVKAWOTKEYNLVRNV 394
DB 361 peycdwqyiglpdqiksvkawotkeynlvrnv 394
RESULT 4
AAR80016
ID AAR80016 standard; Protein; 376 AA.

```

```

XX AAR80016;
XX AC 01-MAY-1996 (first entry)
XX DT Marmoset alpha-1,3-galactosyltransferase.
XX DE Marmoset; alpha-1,3-galactosyltransferase; immune response; glycoprotein;
XX KW alpha-galactosyl epitope; cell membrane; virus; phagocytosis; tumour;
XX KW antigen processing; leukemia; lymphoma; myeloma; melanoma; carcinoma;
XX KW sarcoma; vaccine; opsonisation; glycoprotein; antibody; anti-Gal.
XX OS Callithrix jacchus.
XX PN WO9524924-A1.
XX PD 21-SEP-1995.
XX PF 13-MAR-1995; 95WO-US03156.
XX PR 15-MAR-1994; 94US-0213200.
XX PA (UYHA-) UNIV HAHNEMANN & MEDICAL COLLEGE PENNSYL.
XX PI Gallili U, Replik PM;
XX DR WPI; 1995-336816/43.
XX DR N-PSDB; AAT04522.
XX PT Association of an alpha-galactosyl epitope with a tumour or viral
XX PT antigen - is administered to anti-Gal synthesising animals to induce
XX PS an immune response
XX PS Disclosure; Fig 9; 85pp; English.
XX CC The amino acid sequence of the marmoset alpha-1,3-galactosyltransferase.
XX CC The enzyme can be used in methods of enhancing an immune response by
XX CC associating the alpha-galactosyl epitope with a cell membrane or viral
XX CC glycoprotein. The alpha-galactosyl epitope enhances phagocytosis and
XX CC subsequent processing of the antigen. The method is useful in the
XX CC treatment of tumours e.g. leukemia, lymphoma, myeloma, melanoma,
XX CC carcinoma and sarcoma, or for the generation of viral vaccines by
XX CC opsonising a viral glycoprotein. The alpha-galactosyl epitope enhances
XX CC recognition of the antigen in an animal that synthesises the naturally
XX CC occurring antibody - anti-Gal.
XX SQ Sequence 376 AA;
Query Match 72.1%; Score 1539.5; DB 16; Length 376;
Best Local Similarity 72.6%; Pred. No. 1.2e-146;
Matches 273; Conservative 43; Mismatches 43; Indels 17; Gaps 2;
QY 36 MNVKGKVVILLMLIVSTVVVFWEYVN-----RPEVGENRQKDWMPSPFKN 83
DB 1 mnvkgkvillmlivstvvvfweyinspegflwyhsknpdevdssaqkdwfpfgvfn 60
QY 84 GTHSYQE-----DNVEGRREKGRNGDRIEPPQLWDFWPNKRPDLTVTPWKAPVWEGT 138
DB 61 ghnyqqeecedtkegreeqkddtdtelrlwdfnpkkrpemtvtqkavvvwgt 120
QY 139 YDTALLEKYATOKLTVGLTVFAVGKYIEHYLEDLESADMYFMVGHVRVIFVYVIMDDTSR 198
DB 121 ynkaillenyyakqkltvgtlvfalgrylehylesfvtsanryfmvghkvifvymvddvsk 180
QY 199 MPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHLIAHIQHEVDLFLCNDVQVDFOD 258
DB 181 pfelgplrsfkfvefkpekrwdqldismrmktigehlliahiqhevdfllcndvqvfq 240
QY 259 NFGVETLQGLVAQLQAWMYKASPEKFTYERRELSAAIIPFGEQDFYHAAIFGGTPTHIL 318
DB 241 hfvetlqgsvaqlqawmykadpddftyerrksaaylpfggdfyyhaaifggtpqlvl 300

```



```
QY 94 EGRREGRNDRTEEPOLWDFNPKNRPOVLTVTPWKAPIVWEGTYDTALLEKYATQKL 153
Db 60 gnekeqrkednrgelp-lvdwfnpekrpevvtitrwkapvwwegtyrnarvldnyyakqi 118
QY 154 TVGLTVFAVCKYIEHYLEDFLESADMYFMVGHVIFVYVIMDDTSRMPVVLNPLHLSLOVF 213
Db 119 tvgltvfavgyryehyleeflisantyfmvghkvifymvddlsrmpllieglprskvf 178
QY 214 EIRSEKRWQDISMRMKTIGEHLAHIQHEVDLFCHMDVDQVFDQNFVETLGLVAQLQ 273
Db 179 eiksekrwqdismrmtktigehllahiqhevdficldvdqvfgnfgvetlqgsvaqlq 238
QY 274 AMYKASPEKFTYERRELSAAYIPFGGDDFYHAAIFGGTPTTHILNLTRECFKGILODKK 333
Db 239 awwykapdeftyerrkesaayipfgggdfyyhaaifggtpqvlntqcfkglldke 298
QY 334 HDIEAQWHDHSHLNKYPFLFNKPKILSPCYCWDYQIGLPSDIKSVKVAWQTKKEYNLVRNN 393
Db 299 ndieaewhdeshlnkyfllnkpkilspcywdyhgmsvdirivkiawdkkeynlvrnn 358
QY 394 V 394
Db 359 i 359
RESULT 7
ID AAR90573 standard; Protein; 359 AA.
XX AAR90573;
XX
DT 08-APR-1996 (first entry)
DE Pig alpha(1.3)-galactosyltransferase.
XX
KW Alpha(1.3)galactosyltransferase; xenograft hyperacute rejection;
KW transplantation; galactose alpha(1.3) galactose.
XX
OS Sus scrofa.
XX
PN WQ9534202-A1.
XX
PD 21-DEC-1995.
XX
PF 14-JUN-1995; 95WO-US07554.
XX
PR 21-JUL-1994; 94US-0278282.
PR 13-JUN-1994; 94US-0260201.
XX
PA (ALEX-) ALEXION PHARM INC.
PA (AUST-) AUSTIN RES INST.
XX
PI Fodor WL, McKenzie IFC, Rother RP, Sandrin MS, Squinto SP;
XX
DR WPI; 1996-049326/05.
DR N-PSDB; AAT12242.
XX
PT Redn. of rejection of xenogeneic cells following transplantation
PT by introducing a vector expressing fucosyl:transferase into the
PT cells
XX
PS Example 2; Page 52-54; 69pp; English.
XX
CC Pig alpha(1.3)-galactosyltransferase (AAR90573) was expressed in
CC monkey COS cells following transfection of the cells with
CC vector pGT which contains an insert including the encoding
CC cDNA (AAT12242). Co-transfection of these cells with vector pHT
CC encoding human H-transferase (AAR90572) resulted in a reduction in
CC the levels of galactose alpha(1.3) galactose epitopes expressed
CC by the cells.
XX
SQ Sequence 359 AA;
```

```
Query Match 71.4%; Score 1525; DB 17; Length 359;
Best Local Similarity 75.1%; Pred. No. 3.3e-145;
Matches 271; Conservative 48; Mismatches 38; Indels 4; Gaps 3;
QY 36 MNVKGKAVILLMLIVSTVVVFWFVYVIRIPVGENRQKQDWFFSWFKNGTHSY--QEDNV 93
Db 1 mnvkgkrvvlmlivstvmvfwfyrinrnpvgssa-qrgwfwfswfnngthsyheeda 59
QY 94 EGRREGRNDRTEEPOLWDFNPKNRPOVLTVTPWKAPIVWEGTYDTALLEKYATQKL 153
Db 60 gnekeqrkednrgelp-lvdwfnpekrpevvtitrwkapvwwegtyrnarvldnyyakqi 118
QY 154 TVGLTVFAVCKYIEHYLEDFLESADMYFMVGHVIFVYVIMDDTSRMPVVLNPLHLSLOVF 213
Db 119 tvgltvfavgyryehyleeflisantyfmvghkvifymvddlsrmpllieglprskvf 178
QY 214 EIRSEKRWQDISMRMKTIGEHLAHIQHEVDLFCHMDVDQVFDQNFVETLGLVAQLQ 273
Db 179 eiksekrwqdismrmtktigehllahiqhevdficldvdqvfgnfgvetlqgsvaqlq 238
QY 274 AMYKASPEKFTYERRELSAAYIPFGGDDFYHAAIFGGTPTTHILNLTRECFKGILODKK 333
Db 239 awwykapdeftyerrkesaayipfgggdfyyhaaifggtpqvlntqcfkglldke 298
QY 334 HDIEAQWHDHSHLNKYPFLFNKPKILSPCYCWDYQIGLPSDIKSVKVAWQTKKEYNLVRNN 393
Db 299 ndieaewhdeshlnkyfllnkpkilspcywdyhgmsvdirivkiawdkkeynlvrnn 358
QY 394 V 394
Db 359 i 359
RESULT 8
AAR49686
ID AAR49686 standard; Protein; 375 AA.
XX AAR49686;
XX
DT 10-NOV-1998 (first entry)
DE Porcine alpha-1,3-galactosyl transferase isoform 1.
XX
KW Isoform; porcine; enzyme; alpha-1,3-galactosyl transferase; galactose;
KW sugar; N-acetyllactosamine; glycoprotein; glycolipid; antibody; pig;
KW graft tissue rejection; organ transplantation; xenotransplant.
XX
OS Sus scrofa.
XX
PN PR2751346-A1.
XX
PD 23-JAN-1998.
XX
PF 19-JUL-1996; 96PR-0009077.
XX
PR 19-JUL-1996; 96PR-0009077.
XX
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Pourcel C, Soullilou JP, Vanhove B;
XX
DR WPI; 1998-112876/11.
DR N-PSDB; AAV49453.
XX
PT Transgenic non-human donors of organs for human recipients -
PT containing DNA encoding antibodies that inhibit graft rejection
XX
PS Claim 4; Page 32-34; 71pp; French.
XX
CC This sequence represents isoform 1 of the porcine enzyme
CC alpha-1,3-galactosyl transferase (alpha-1,3-GT). The enzyme catalyses
```

CC the attachment of a galactose sugar molecule on the N-acetyllactosamine  
 CC molecule found on surface glycoproteins and glycolipids. These sugar  
 CC molecules are partly responsible for raising anti-graft antibodies, which  
 CC lead to graft tissue rejection. The invention relates to a method of  
 CC inhibiting the graft rejection mechanism by introducing the sequence  
 CC encoding an antibody targeted to alpha-1,3-GT into the cells of animal,  
 CC especially a pig, from whom organs may be used for xenotransplants.  
 CC Neutralisation of the alpha-1,3-GT leads to tissues or organs lacking  
 CC the galactose on the glycoproteins and glycolipids, thus preventing  
 CC induction of the rejection response.  
 XX  
 SQ Sequence 375 AA;

Query Match 70.8%; Score 1513; DB 19; Length 375;  
 Best Local Similarity 72.5%; Pred. No. 5.8e-144;  
 Matches 272; Conservative 48; Mismatches 39; Indels 16; Gaps 4;  
 QY 34 KIMVKGKIVILLMLIVSTVVVWFVEYVN-----RIPEVGENRQKDWPFPSWF 81  
 DB 3 KIMVKGKIVILLMLIVSTVVVWFVEYVN-----RIPEVGENRQKDWPFPSWF 61  
 QY 82 KNGTHSY--QEDNVEGRREKGRNGDRTEEPQWDFNPKNRPDLTVTPWKPAPIWEGTY 139  
 DB 62 nngthsyheedaignekeqrkednrgelp-lvdwfnpekrpevvlttrwkapvwegty 120  
 QY 140 DTALLEKYYATQKLTVGLTVFVAVGKYIEHYLEDFLESADMYFMVGHRYFYVMIDDTSRM 199  
 DB 121 nradlmyyakkltvgltvavgrylehyieeflisantyfmvghkvifvlymvdldsrml 180  
 QY 200 PVVHLNPLHSLOVFEIRSEKRWQDISMMRMKTIQEHILAHQHEVDLFCMDVDVQDN 259  
 DB 181 plielgplrsfkveiksekrgwdismmrmtkigeihlahiqhevdflfcmvddvqfnq 240  
 QY 260 FGVTETGLQVAQLQAWMYKASPEKFTYERRELSSAAYIPFGGDFYHAAIFGGTPTHTLN 319  
 DB 241 fgvetlgqsvaqlqawmykahpdeftyerrkesaaylpgggdfyhaaifggcptqvin 300  
 QY 320 LTRECFKIGLODKKHIDEAQWHDHSHLNKYLFPNKPTKILSPYCWYQIGLPSDIKSVK 379  
 DB 301 itqecfgilqdkendieawhdeshlnkyflinkptkilspeycwdyhgmsvdirivk 360  
 QY 380 VAWQTKENLVNRNV 394  
 DB 361 iawqtkenlvnrnni 375

RESULT 9  
 AAR85082  
 ID AAR85082 standard; Protein: 371 AA.  
 XX

AC AAR85082;  
 XX  
 XX  
 DT 04-JUN-1996 (first entry)  
 XX

DE Porcine alpha (1,3) galactosyltransferase.  
 XX

KW Transgenic; swine; porcine; alpha (1,3) galactosyltransferase;  
 KW antisense; ribozyme; Gal-alpha-1,3-Gal-beta-1-4GlcNAc; epitope;  
 KW terminal; xenogenic; transplant; rejection; gene therapy; pig.  
 XX  
 OS Sus scrofa.

XX WO9528412-A1.  
 XX  
 XX 26-OCT-1995.  
 XX  
 XX 31-MAR-1995; 95WO-US03940.  
 XX  
 XX 13-APR-1994; 94US-0228933.  
 XX  
 XX (BIOT-) BIOTRANSPLANT INC.  
 PA (GEHO ) GEN HOSPITAL CORP.

(CHIL-) INST CHILD HEALTH.

XX Baetscher MW, Gustafsson KT, Sachs DH;

XX WPI: 1995-373759/48.

DR N-PSDB: AAT02892.

XX Novel transgenic alpha (1,3) galactosyltransferase negative swine  
 PT - used to produce rejection resistant cells for xenogenic  
 PT transplantation

XX Claim 11; Pages 35-37; 56pp; English.

XX Transgenic swine in which the normal expression of the alpha (1,3)  
 CC galactosyltransferase (AGT) AAR85082 is prevented, are prep. by  
 CC inhibiting the expression of the AGT gene AAT02892 using antisense  
 CC oligonucleotides or ribozyme inactivators in a pluripotent porcine  
 CC embryonic stem cell. It is then inserted into a porcine oocyte  
 CC (from which the pronuclear material has been removed), which is  
 CC itself grown to produce the transgenic swine. Swine which do not  
 CC express AGT will not produce carbohydrate moieties contg. the  
 CC distinctive terminal Gal-alpha-1,3-Gal-beta-1-4GlcNAc epitope,  
 CC which is a significant factor in xenogenic (esp. human) transplant  
 CC rejection of swine grafts. Therefore the swine cells produced in  
 CC the AGT negative transgenic swine are xenogenic transplant  
 CC rejection resistant, and can therefore be used by a transplant  
 CC recipient, or to provide gene therapy.

XX Sequence 371 AA;

Query Match 70.2%; Score 1500; DB 16; Length 371;  
 Best Local Similarity 72.4%; Pred. No. 1.2e-142;  
 Matches 270; Conservative 47; Mismatches 40; Indels 16; Gaps 4;

QY 36 MNVKGKIVILLMLIVSTVVVWFVEYVN-----RIPEVGENRQKDWPFPSWFKN 83  
 DB 1 mnvkgkivillmlivstvmvfweyinspegslfwyqsknpevgssa-grgwfwpswfn 59  
 QY 84 GTHSY--QEDNVEGRREKGRNGDRTEEPQWDFNPKNRPDLTVTPWKPAPIWEGTYDT 141  
 DB 60 gthsyheedaignekeqrkednrgelp-lvdwfnpekrpevvlttrwkapvwegty 118  
 QY 142 ALLEKYYATQKLTVGLTVFVAVGKYIEHYLEDFLESADMYFMVGHRYFYVMIDDTSRMPV 201  
 DB 119 avldnyyakqkltvltvavgrylehyieeflisantyfmvghkvifvlymvdldsrml 178  
 QY 202 VHLNPLHSLOVFEIRSEKRWQDISMMRMKTIQEHILAHQHEVDLFCMDVDVQDNFG 261  
 DB 179 ielgplrsfkveiksekrgwdismmrmtkigeihlahiqhevdflfcmvddvqfnq 238  
 QY 262 VETLGQVAQLQAWMYKASPEKFTYERRELSSAAYIPFGGDFYHAAIFGGTPTHTLNLT 321  
 DB 239 vetlgqsvaqlqawmykahpdeftyerrkesaaylpgggdfyhaaifggcptqvin 298  
 QY 322 RECFKIGLODKKHIDEAQWHDHSHLNKYLFPNKPTKILSPYCWYQIGLPSDIKSVKVA 381  
 DB 299 qecfkgilqdkendieawhdeshlnkyflinkptkilspeycwdyhgmsvdirivk 358  
 QY 382 WQTKENLVNRNV 394  
 DB 359 wqtkenlvnrnni 371

RESULT 10  
 AAR76777  
 ID AAR76777 standard; Protein: 371 AA.  
 XX

AC AAR76777;  
 XX

XX 11-DEC-1995 (first entry)  
 XX  
 DE Pig alpha-1,3-galactosyltransferase.

XX Alpha-1,3-galactosyltransferase; alpha-1,3-GALT; transgenic animal;  
KW pig; hyperacute rejection; xenotransplantation; donor organ;  
KW allograft rejection; Gal epitope; gene disruption;  
KW homologous recombination; knock-out.  
XX  
OS Sus scrofa.  
XX W09520661-A1.  
PN  
XX  
PD  
XX  
XX 03-AUG-1995.  
XX  
PF 27-JAN-1995; 95WO-IB00088.  
XX  
XX 26-JAN-1995; 95US-0188607.  
PR 27-JAN-1994; 94US-0188607.  
XX  
XX (BRES-) BRESATEC LTD.  
PA (SVIN-) ST VINCENT'S HOSPITAL MELBOURNE LTD.  
XX  
XX Crawford RJ, Dapice AJF, Pearse MJ, Rathjen PD;  
PI Robbins AJ;  
XX  
XX WPI: 1995-275446/36.  
DR N-PSDB; AAQ93077.  
XX  
XX New alpha-1,3-galactosyltransferase and leukaemia inhibitor factor  
PT - corresp. DNA and nucleic acid constructs for inactivating the  
PT transferase gene; for eliminating hyperacute region in human  
PT transplants  
XX  
XX Claim 3; Fig.5; 184pp; English.  
PS  
XX  
XX cDNA encoding porcine alpha-1,3-GALT was generated from liver RNA  
CC using primers based on conserved regions of the mouse and cattle alpha-  
CC 1,3-GALT genes. Potential sites to interrupt the alpha-1,3-GALT gene  
CC (via homologous recombination) were identified in exons 4, 7, 8 and 9.  
CC Such inactivation allows the breeding of 'knock-out' animals, e.g.  
CC pigs suitable as donors of organs to overcome hyperacute rejection  
CC problems in human xenotransplantation.  
XX  
XX Sequence 371 AA;  
XX  
XX Query Match 69.9%; Score 1493; DB 16; Length 371;  
Best Local Similarity 71.8%; Pred. No. 5.9e-142;  
Matches 268; Conservative 49; Mismatches 40; Indels 16; Gaps 4;  
QY 36 MNVKGKVVILLMLIVSTVVVFWEYVN-----RIPEVGENRWQKDMWFPSPKFN 83  
DB 1 mnvkgrrvllstvmvfweyinspegslfwlqsknpevgssa-qrgwwfswfn 59  
QY 84 GTHSY--QEDNVGRRKGRNGRIEPEQLWDFNPKRPDVLTVTPWKAPIWEGTYDT 141  
DB 60 gthsyheedaagnekeqrkednrgelp-lvdwfnpekrpevvltirwkapvwegtnr 118  
QY 142 ALLEKRYATQKLTGVTFAVGKVIIEHYLEDLESADMYFMVGHVRVIFVIMDDTSRMPV 201  
DB 119 avldnynakqkicvgtfavgryiehyleeiflisanfymvghkvifylmvdiddisrmp 178  
QY 202 VHNLPHLSQVFIRESKRWQDISMMRMKTIGEHILAHQHEVDLFPCMDVDQVFDNFG 261  
DB 179 ielgprsfkfvfkaektwqdismmrmktigehilahiheqvefdlfcmdvdqvfqnfg 238  
QY 262 VETLGOLVAQLOAWKAKSPEKTYERRELSAAYIPFGEGDFYHAAIFGCTTHILNLT 321  
DB 239 vecigsvaqlqawykahpdefyerrkesaayipfgqgdfyhaaifggtptqvlnt 298  
QY 322 RECFKGLQDKKHDIEAQWHDESHNLKRYLFNKPRTKILSPEYCDWYQIGLPSDKSVKVA 381  
DB 299 qecfkglqdkendieaewhdehnlkylilnkptkilspeycwdyhgmsvdirivkia 358  
QY 382 WQKEYNLVRNVV 394

DB 359 wqkkeynlvrnni 371  
RESULT 11  
AAW49688  
ID AAW49688 standard; Protein; 354 AA.  
XX  
AC AAW49688;  
XX  
DT 10-NOV-1998 (first entry)  
XX  
DE Porcine alpha-1,3-galactosyl transferase isoform 3.  
XX  
XX Isoform; porcine; enzyme; alpha-1,3-galactosyl transferase; galactose;  
KW sugar; N-acetyllactosamine; glycoprotein; glycolipid; antibody; pig;  
KW graft tissue rejection; organ transplantation; xenotransplant.  
XX  
OS Sus scrofa.  
XX  
XX FR2751346-A1.  
PN  
XX 23-JAN-1998.  
PD  
XX 19-JUL-1996; 96FR-0009077.  
PF  
XX 19-JUL-1996; 96FR-0009077.  
PR  
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
PA  
XX Pourcel C, Soullilou JP, Vanhove B;  
PI  
XX WPI: 1998-112876/11.  
DR N-PSDB; AAV49455.  
XX  
XX Transgenic non-human donors of organs for human recipients -  
PT containing DNA encoding antibodies that inhibit graft rejection  
XX  
XX Claim 4; Page 39-41; 71pp; French.  
PS  
XX  
XX This sequence represents isoform 3 of the porcine enzyme  
CC alpha-1,3-galactosyl transferase (alpha-1,3-GT). The enzyme catalyses  
CC the attachment of a galactose sugar molecule on the N-acetyllactosamine  
CC moiety found on surface glycoproteins and glycolipids. These sugar  
CC molecules are partly responsible for raising anti-graft antibodies, which  
CC lead to graft tissue rejection. The invention relates to a method of  
CC inhibiting the graft rejection mechanism by introducing the sequence  
CC encoding an antibody targeted to alpha-1,3-GT into the cells of animal,  
CC especially a pig, from whom organs may be used for xenotransplants.  
CC Neutralisation of the alpha-1,3-GT leads to tissues or organs lacking  
CC the galactose on the glycoproteins and glycolipids, thus preventing  
CC induction of the rejection response.  
XX  
XX Sequence 354 AA;  
XX  
XX Query Match 67.0%; Score 1431.5; DB 19; Length 354;  
Best Local Similarity 72.2%; Pred. No. 8.7e-136;  
Matches 262; Conservative 48; Mismatches 40; Indels 13; Gaps 5;  
QY 34 KIMNVKGKVVILLMLIVSTVVVFWEYVNRIPEVGENRWQKDMWFPSPKFNTHSY--QED 91  
DB 3 klmnvkgrrvllstvmvfweyins-pe-gslfw-----lyqskthsheed 52  
QY 92 NVGRRKGRNGRIEPEQLWDFNPKRPDVLTVTPWKAPIWEGTYDTALLEKRYATQ 151  
DB 53 aignekeqrkednrgelp-lvdwfnpekrpevvltirwkapvwegtnravldnvyak 111  
QY 152 KLTGVTFAVGKVIIEHYLEDLESADMYFMVGHVRVIFVIMDDTSRMPVHNLPHLSQ 211  
DB 112 kicvgtfavgryiehyleeiflisanfymvghkvifylmvdiddisrmpfgrsfk 171  
QY 212 VPEIRSEKRWQDISMMRMKTIGEHILAHQHEVDLFPCMDVDQVFDNFGVETLGOLVAQ 271

```
Db 172 vfeiksekrwgdismmrkktigeihlahighedvflfcmddvqvgfnfgvetlgsaq 231
QY 272 LOAWMYKASPEKFTYERRELSSAAVIFPGEGDYFYHAAIFGCTPTHILNLTRECFKGILOD 331
Db 232 lqawwykahpdefcyerrkesaayipfgqgdfyyhaaifggtctqvinicqectkgllqd 291
QY 332 KKHIDIAQWHDHSHLNKYFLFNKPTKILSPCYCWDYQIGLPSDIKSVKVAQTKKEYNLVR 391
Db 292 kendieaewhdeshlnkylfllnkptkilspeycwdyhgmsvdirivkikawqkkeynlvr 351
QY 392 NNV 394
Db 352 nni 354

RESULT 12
AAW49689
ID AAW49689 standard; Protein; 342 AA.
XX
AC AAW49689;
XX
DT 10-NOV-1998 (first entry)
XX
DE Porcine alpha-1,3-galactosyl transferase isoform 4.
XX
KW Isoform; porcine; enzyme; alpha-1,3-galactosyl transferase; galactose;
KW sugar; N-acetylglucosamine; glycoprotein; glycolipid; antibody; pig;
KW graft tissue rejection; organ transplantation; xenotransplant.
XX
OS Sus scrofa.
XX
PN FR2751346-A1.
XX
PD 23-JAN-1998.
XX
PF 19-JUL-1996; 96FR-0009077.
XX
PR 19-JUL-1996; 96FR-0009077.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Pourcel C, Soullilou JP, Vanhove B;
XX
DR WPI; 1998-112876/11.
DR N-PSDB; AAV49456.
XX
PT Transgenic non-human donors of organs for human recipients -
PT containing DNA encoding antibodies that inhibit graft rejection
XX
PS Claim 4; Page 43-44; 71pp; French.
XX
CC This sequence represents isoform 4 of the porcine enzyme
CC alpha-1,3-galactosyl transferase (alpha-1,3-GT). The enzyme catalyses
CC the attachment of a galactose sugar molecule on the N-acetylglucosamine
CC moiety found on surface glycoproteins and glycolipids. These sugar
CC molecules are partly responsible for raising anti-graft antibodies, which
CC lead to graft tissue rejection. The invention relates to a method of
CC inhibiting the graft rejection mechanism by introducing the sequence
CC encoding an antibody targeted to alpha-1,3-GT into the cells of animal,
CC especially a pig, from whom organs may be used for xenotransplants.
CC Neutralisation of the alpha-1,3-GT leads to tissues or organs lacking
CC the galactose on the glycoproteins and glycolipids, thus preventing
CC induction of the rejection response.
XX
SQ Sequence 342 AA;

Query Match 66.9%; Score 1428.5; DB 19; Length 342;
Best Local Similarity 71.3%; Pred. No. 1.7e-135;
Matches 259; Conservative 45; Mismatches 34; Indels 25; Gaps 3;

QY 34 KIMNVKGVILLMLIVSTVVWFVEYVNRIPVEGENRQKDMWFPSPFKNGTHSY--QED 91
```

```
Db 3 kimmvkgrrvvlsmllvstvmvfweylnr-----thsyheeed 40
QY 92 NVEGRREKGRNGDRIRBEPQOLWDNPNKRNPDVLTVTWKAPIVWEGTYDTALLKYYATO 151
Db 41 aignekeqrkednrgelp-lvdwfnpekrpevctirwkapvwwegtyrnrvlndnyyakq 99
QY 152 KLTVGLTFVAVGKYIEHYLEDLFLESADMYFMVGHVRVIFYVMIDDTSRMPVVLNPLHSLQ 211
Db 100 kitvgltvavgrylehyleeefliisantyemvghkvifylmvddisrmpilelglprsfk 159
QY 212 VFEIRSEKRWQDTSMMRMKKTIGEHLAHIGHEDVDFLCMDVDQVFDQNFQVETLGLQVLAQ 271
Db 160 vfeiksekrwgdismmrkktigeihlahighedvflfcmddvqvgfnfgvetlgsaq 219
QY 272 LOAWMYKASPEKFTYERRELSSAAVIFPGEGDYFYHAAIFGCTPTHILNLTRECFKGILOD 331
Db 220 lqawwykahpdefcyerrkesaayipfgqgdfyyhaaifggtctqvinicqectkgllqd 279
QY 332 KKHIDIAQWHDHSHLNKYFLFNKPTKILSPCYCWDYQIGLPSDIKSVKVAQTKKEYNLVR 391
Db 280 kendieaewhdeshlnkylfllnkptkilspeycwdyhgmsvdirivkikawqkkeynlvr 339
QY 392 NNV 394
Db 340 nni 342

RESULT 13
AAR62507
ID AAR62507 standard; Protein; 313 AA.
XX
AC AAR62507;
XX
DT 26-JUN-1995 (first entry)
XX
DE Galactosyl transferase 3' clone product.
XX
KW Gal-alpha (1,3) galactosyl transferase; xenograft; transplant;
KW rejection.
XX
OS Sus scrofa domestica.
XX
PN WO9421799-A.
XX
PD 29-SEP-1994.
XX
PF 15-MAR-1994; 94WO-AU00126.
XX
PR 16-MAR-1993; 93AU-0007854.
XX
PA (AUST-) AUSTIN RES INST.
XX
PI McKenzie IFC, Sandrin MS;
XX
DR WPI; 1994-317019/39.
DR N-PSDB; AAQ74711.
XX
PT DNA sequences encoding Gal-alpha (1,3)galactosyl transferase -
PT and clones contg. such sequences are used in xenograft therapies
XX
PS Disclosure; Page 31; 50pp; English.
XX
CC The sequence is that of the product of the porcine Gal-alpha (1,3)
CC galactosyl transferase gene which produces a Gal epitope on the
CC surface of porcine cells. This epitope is recognised by antibodies
CC which are responsible for hyperacute rejection of xenotransplanted
CC pig cells, tissues and organs.
CC See also AAR62508.
XX
SQ Sequence 313 AA;
```





